SIAM J. Control Optim.
Vol. 39, No. 4, pp. 1233–1295

A TWOFOLD SPLINE APPROXIMATION FOR FINITE HORIZON LQG CONTROL OF HEREDITARY SYSTEMS*

A. GERMANI†, C. MANES†, and P. PEPE†

Abstract. In this paper an approximation scheme is developed for the solution of the linear quadratic Gaussian (LQG) control on a finite time interval for hereditary systems with multiple noncommensurate delays and distributed delay. The solution here proposed is achieved by means of two approximating subspaces: the first one to approximate the Riccati equation for control and the second one to approximate the filtering equations. Since the approximating subspaces have finite dimension, the resulting equations can be implemented. The convergence of the approximated control law to the optimal one is proved. Simulation results are reported on a wind tunnel model, showing the high performance of the method.

Key words. hereditary systems, linear quadratic Gaussian regulator, infinite dimensional systems, Galerkin spline approximation

AMS subject classifications. 93E11, 93E20, 93E25

PII. S0363012998337461

1. Introduction.

It is well known that the solution of the linear quadratic regulation problem and of the optimal Gaussian filtering problem for linear delay systems is found in terms of infinite dimensional operators [7, 8, 9, 10, 11, 12, 13, 17, 23, 31, 36, 37, 39]. On the other hand, implementation of a control/filtering scheme in this case requires a finite dimensional approximation of such operators.

Although much attention has been devoted to separately developing an approximation theory for the linear quadratic (LQ) regulation [4, 11, 12, 16, 24, 26, 27, 30, 33, 40] and the optimal Gaussian filtering [14, 20] of delay systems, the approximation problem of the overall linear quadratic Gaussian (LQG) regulator has not been conveniently treated in the literature.

The averaging approximation scheme has been used in [24], for both the finite and infinite horizon LQ problem of delay systems, and convergence results are obtained by considering a conjecture, later proved to be true [41], that is the question of whether the sequence of approximating systems gives uniformly exponentially stable systems for sufficiently large indexes if the underlying retarded functional differential equation is stable.

The spline approximation scheme developed in [3] has been applied to the LQ problem of delay systems in [4]. Although numerical simulations show better performance than the averaging scheme, no theoretical convergence results are so far available. In [6], it is proved that the adjoint of the approximate semigroup governing the system does not converge in a strong way to the adjoint of such semigroup. As a consequence, the main hypothesis which guarantees the convergence results in [24] cannot be satisfied, and therefore this spline approximation scheme cannot be safely applied.

A new spline approximation scheme has been developed in [27], for the LQ prob-

*Received by the editors April 15, 1998; accepted for publication (in revised form) July 6, 2000; published electronically December 13, 2000. This work was supported by ASI (Agenzia Spaziale Italiana). Preliminary results have been published in [19, 21].
http://www.siam.org/journals/sicon/39-4/33746.html

†Dipartimento di Ingegneria Elettrica, Università degli Studi dell'Aquila, 67040 Poggio di Roio, L'Aquila, Italy (germani@ing.univaq.it, manes@ing.univaq.it, pepe@ing.univaq.it).

lem of delay systems with any number of pure delay terms, assuming the absolute continuity of the kernel in the distributed delay integral. Theoretical convergence results are obtained in the finite horizon case, as this approximation scheme does not guarantee the uniform exponential stability of the approximate semigroups. However, it is proved in [28, 29] that, in the case of commensurate delays and without distributed delay, a weaker condition is sufficient to obtain the strong convergence of the approximated LQ algebraic Riccati equation solution. The authors call this condition *uniform output stability*. It is proved that the spline approximation scheme developed in [27] does satisfy this condition, so that the above convergence result is available for the infinite horizon case. But, as pointed out by Morris on page 9 of paper [36], the convergence properties of this approximation scheme are not sufficient to ensure convergence of the closed loop response.

In [40], a piecewise linear approximation theory has been developed for the finite and infinite horizon LQ of general delay systems. Theoretical convergence results are obtained both in the finite and infinite horizon cases, as the condition of uniform exponential stability is verified.

In [32] error estimates are established for the approximation of delay systems by means of the averaging scheme. In [26] a scheme using first order splines is developed satisfying the uniform exponential stability condition, and error estimates are established too, as is done in [32] for the averaging scheme. Such a scheme uses the classic averaging subspace of piecewise constant functions to define the approximated system equation, but defines the approximated infinitesimal generator in that subspace not in the usual averaging methodology but by using an inverse projector from such subspace to the subspace built up using splines. Such a scheme, which is a mixed averaging spline one, is used in [26] for the infinite horizon LQ problem of general hereditary systems.

The matter of uniform exponential stability for spline approximation schemes has been investigated in [15], in the scalar open loop case. There the real eigenvalue (unique if the coefficient on delay term is positive, in the hereditary equation) of the infinitesimal generator of the semigroup governing the system is used, in order to define a particular inner product, by which Galerkin spline approximations [3] preserve the uniform exponential stability of the approximated semigroups. How this can be applied to optimal multivariables regulator problems is an open and interesting question.

In the synthesis of approximate optimal controllers developed by all above approximation schemes [4, 24, 26, 27, 40] it is assumed that the system state is completely accessible. Moreover, the approximated control input is generated by a finite rank feedback operator applied to the true state in the delay time interval. From an engineering point of view, the resulting controller is still infinite dimensional and therefore not directly implementable.

The synthesis of finite dimensional dynamic output feedback compensators for hereditary systems in a deterministic setting is considered in paper [25, section 4.2]. The proposed controller is composed of an observer and of a feedback control law from the observed state. Both the gains, for the finite dimensional observer and control, are obtained by approximating the solutions of two algebraic Riccati equations. The resulting controller resembles the solution of an LQG problem, although no reference to an optimal stochastic control problem is made in the paper. The main tool is the use of the averaging approximation scheme [2, 24] and the main result is the stability of the overall closed loop system.

In [35] the same problem is investigated with reference to a general class of deterministic distributed systems.

An approximation theory that provides an implementable scheme for the filtering problem of systems evolving on Hilbert spaces has been studied in [14, 18, 20, 22]. This theory has been successfully applied to delay systems.

In the literature the case with one pure delay term is usually completely reported [2, 24, 28, 29, 40] and the general case with multiple noncommensurate delays is usually just briefly indicated. However, the extension of all results to the general case is not straightforward [2, 3, 24, 40] or even unfeasible [28, 29].

As a final point of this bibliographic review, we must stress the existence of a large amount of spline approximation schemes [3, 4, 26, 27, 40] for the deterministic optimal quadratic state regulator (LQ problem), where the control gain operator is approximated by approximating the relevant Riccati equation. In principle, the same approximation schemes could be adapted for approximating the covariance operator defined by the solution of the dual Riccati equation, and the Kalman filter equation that solves the LQG problem in the stochastic setting. On the other hand, the applicability of such schemes to the case of stochastic delay systems with partial noisy state observations is not a trivial question and it has not been investigated up to now, and the main problem of proving the convergence remains unsolved.

The control problem with partial state observation has been treated in literature employing the averaging scheme in a deterministic setting [25]. On the other hand, a known result [4] is the superiority of spline approximation schemes with respect to averaging ones, with respect to numerical convergence rate.

On the basis of these considerations the aim of this paper is to define a finite dimensional scheme that approximates the solution of the finite horizon linear quadratic Gaussian control problem for stochastic delay systems with partial observations. The resulting implementable scheme has the following features:

(i) the optimal closed loop response of the LQG problem can be approximated with arbitrarily small error;
(ii) the scheme can be applied also in the LQ problem;
(iii) the approximation method is based on splines and not on averaging;
(iv) the matrices that implement the approximation of the optimal filter-controller scheme are easily parametrized as a function of the approximation order and can be easily computed;
(v) the scheme allows one to deal with general hereditary systems, that is, with multiple noncommensurate delays and distributed delay;
(vi) simultaneous approximation of a semigroup and of its adjoint is not required, so that problems arising from nondensity of the intersection of the respective generator domains are avoided;
(vii) the scheme allows a quite natural extension to be used for the solution of the infinite horizon LQG problem;
(viii) the scheme has nice numerical properties, in that it shows good performances even with a low finite dimensional approximation order;
(ix) the scheme allows one to get a faster convergence of the approximation by increasing the order of the spline degree.

Of course for most of the above-mentioned points, the scientific literature offers effective algorithms. Nevertheless, the problem of considering all these issues at the same time remains an interesting point.

The paper is organized as follows. In section 2 stochastic hereditary systems are written in state-space form and the infinitesimal generator of the adjoint of the semigroup that governs the system is studied. It is proved that such an operator has a deeply different structure if a weighted inner product is used instead of the usual one. In section 3 the finite horizon LQG is presented, and theorems for a suitable approximation scheme are proved. In section 4 an approximation scheme which satisfies hypotheses of section 3 is described for the general case. In section 5 matrices which represent finite dimensional linear operators are calculated to implement the method. In section 6 the infinite horizon case is addressed. In section 7 simulation results are reported, showing the effectiveness of the proposed method. Section 8 contains the conclusions.

2. Stochastic delay systems. In this paper we deal with the class of those dynamical systems that in technical literature are generally known as *linear delay systems*, sometimes also called *hereditary systems*. When state and observation noise are present, these are described, for $t \geq 0$, by stochastic equations of the type

$$\begin{aligned} \dot{\boldsymbol{z}}(t) &= \boldsymbol{A}_0\boldsymbol{z}(t) + \sum_{h=1}^{\delta}\boldsymbol{A}_h\boldsymbol{z}(t-r_h) \\ &\quad + \int_{-r}^{0}\boldsymbol{A}_{01}(\vartheta)\boldsymbol{z}(t+\vartheta)d\vartheta + \boldsymbol{B}_0\boldsymbol{u}(t) + \boldsymbol{F}_0\boldsymbol{\omega}(t), \\ \boldsymbol{y}(t) &= \boldsymbol{C}_0\boldsymbol{z}(t) + \boldsymbol{G}\boldsymbol{\omega}(t) \end{aligned} \tag{2.1}$$

with $\boldsymbol{z}(t) \in \mathbb{R}^N$, $\boldsymbol{u}(t) \in \mathbb{R}^p$, $\boldsymbol{y}(t) \in \mathbb{R}^q$, $\boldsymbol{\omega}(t) \in \mathbb{R}^s$, $r_\delta = r > r_{\delta-1} > \cdots r_1 > r_0 = 0$, $\boldsymbol{A}_h \in \mathbb{R}^{N\times N}$, $\boldsymbol{A}_{01} \in L_2([-r,0];\mathbb{R}^{N\times N})$, $\boldsymbol{B}_0 \in \mathbb{R}^{N\times p}$, $\boldsymbol{C}_0 \in \mathbb{R}^{q\times N}$, $\boldsymbol{G} \in \mathbb{R}^{q\times s}$, $\boldsymbol{F}_0 \in \mathbb{R}^{N\times s}$.

The noise $\boldsymbol{\omega}$ belongs to the Hilbert space $L_2([0,t_f];\mathbb{R}^s)$ equipped with the standard Gaussian cylinder measure (this corresponds to model $\boldsymbol{\omega}$ as a white-noise process [1]). Independence of state and observation noises is assumed, that is, $\boldsymbol{F}_0\boldsymbol{G}^{\mathrm{T}} = \boldsymbol{0}$ and, without loss of generality, $\boldsymbol{G}\boldsymbol{G}^{\mathrm{T}} = \boldsymbol{I}_q$, where $\boldsymbol{I}_q$ denotes the identity matrix in $\mathbb{R}^{q\times q}$.

The variable $\boldsymbol{z}$ in the interval $[-r,0]$ is assumed to be generated as follows:

$$\boldsymbol{z}(\vartheta) = \bar{\boldsymbol{z}}(\vartheta) + \int_{-r}^{0} k(\vartheta,\tau)\bar{\boldsymbol{\omega}}(\tau)d\tau, \quad \vartheta \in [-r,0], \tag{2.2}$$

where $\bar{\boldsymbol{z}}$ is absolutely continuous with derivative in $L_2([-r,0];\mathbb{R}^N)$ and the process $\bar{\boldsymbol{\omega}}$, independent of $\boldsymbol{\omega}$, belongs to the Hilbert space $L_2([-r,0];\mathbb{R}^{\bar{s}})$ equipped with the standard Gaussian cylinder measure, and the kernel $k(\vartheta,\tau)$ is integrable for $\tau \in [-r,0]$.

As is well known, system (2.1) can be rewritten in state-space form in the Hilbert space $\boldsymbol{M}_2 = \mathbb{R}^N \times L_2([-r,0];\mathbb{R}^N)$, endowed with the following weighted inner product [3]:

$$\left(\begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix}, \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{bmatrix} \right)_{\boldsymbol{M}_2} = \boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{y}_0 + \int_{-r}^{0}\boldsymbol{x}_1^{\mathrm{T}}(\vartheta)\boldsymbol{y}_1(\vartheta)g(\vartheta)d\vartheta, \tag{2.3}$$

where $g(\vartheta)$ is the piecewise constant nondecreasing function defined as

$$g(\vartheta) = \chi_{[-r_\delta,-r_{\delta-1}]}(\vartheta) + \sum_{j=1}^{\delta-1}(\delta-j+1)\chi_{(-r_j,-r_{j-1}]}(\vartheta), \tag{2.4}$$

where χ_S denotes the characteristic function of the interval S.

Here and in the following the standard assumption is made that summations vanish when the upper limit is smaller than the lower one (e.g., $\delta = 1$ in (2.4)).

In this paper, for the sake of brevity and whenever it does not cause confusion, the space $L_2([-r,0];\mathbb{R}^N)$ will be simply indicated as L_2. In the same way C^k will denote the space $C^k([-r,0];\mathbb{R}^N)$ of functions with values in $\mathbb{R}^N$ that have continuous derivatives until order k, while the symbol $W^{1,2}$ will indicate the space of absolutely continuous functions from $[-r,0]$ in $\mathbb{R}^N$, with derivative in L_2.

In $\boldsymbol{M}_2$ the system (2.1), (2.2) assumes the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{F}\boldsymbol{\omega}(t), \quad \boldsymbol{x}(0) = \begin{bmatrix} \bar{\boldsymbol{z}}(0) \\ \bar{\boldsymbol{z}} \end{bmatrix} + \begin{bmatrix} \mathcal{L}_0 \\ \mathcal{L}_1 \end{bmatrix} \bar{\boldsymbol{\omega}}, \tag{2.5}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{G}\boldsymbol{\omega}(t), \tag{2.6}$$

where $\boldsymbol{A} : \mathcal{D}(\boldsymbol{A}) \mapsto \boldsymbol{M}_2$ is defined as

$$\boldsymbol{A}\begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_0\boldsymbol{x}_0 + \sum_{h=1}^{\delta} \boldsymbol{A}_h\boldsymbol{x}_1(-r_h) + \int_{-r}^{0} \boldsymbol{A}_{01}(\vartheta)\boldsymbol{x}_1(\vartheta)d\vartheta \\ \frac{d}{d\vartheta}\boldsymbol{x}_1 \end{bmatrix} \tag{2.7}$$

with domain

$$\mathcal{D}(\boldsymbol{A}) = \left\{ \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix} \middle| \begin{array}{c} \boldsymbol{x}_0 \in \mathbb{R}^N \\ \boldsymbol{x}_1 \in W^{1,2} \end{array} \quad \boldsymbol{x}_0 = \boldsymbol{x}_1(0) \right\}, \tag{2.8}$$

and the linear operators $\boldsymbol{B}, \boldsymbol{C}, \boldsymbol{F}$ are defined as

$$\boldsymbol{B} : \mathbb{R}^p \mapsto \boldsymbol{M}_2, \qquad \boldsymbol{B}\boldsymbol{u}(t) = \begin{bmatrix} \boldsymbol{B}_0\boldsymbol{u}(t) \\ \boldsymbol{0} \end{bmatrix}, \tag{2.9}$$

$$\boldsymbol{C} : \boldsymbol{M}_2 \mapsto \mathbb{R}^q, \qquad \boldsymbol{C}\begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix} = \boldsymbol{C}_0\boldsymbol{x}_0, \tag{2.10}$$

$$\boldsymbol{F} : \mathbb{R}^s \mapsto \boldsymbol{M}_2, \qquad \boldsymbol{F}\boldsymbol{\omega}(t) = \begin{bmatrix} \boldsymbol{F}_0 \ \boldsymbol{\omega}(t) \\ \boldsymbol{0} \end{bmatrix}. \tag{2.11}$$

The Hilbert–Schmidt operator $\mathcal{L} = \left[\begin{smallmatrix} \mathcal{L}_0 \\ \mathcal{L}_1 \end{smallmatrix}\right]$, which defines the stochastic initial state $\boldsymbol{x}(0)$, derives from definition (2.2) and is defined as follows:

$$\begin{aligned} &\mathcal{L}_0 : \ L_2([-r,0];\mathbb{R}^{\bar{s}}) \mapsto \mathbb{R}^N; \qquad &\mathcal{L}_0\boldsymbol{\omega} = \int_{-r}^{0} k(0,\tau)\bar{\boldsymbol{\omega}}(\tau)d\tau, \\ &\mathcal{L}_1 : \ L_2([-r,0];\mathbb{R}^{\bar{s}}) \mapsto W^{1,2}; \qquad &\mathcal{L}_1\boldsymbol{\omega}(\vartheta) = \int_{-r}^{0} k(\vartheta,\tau)\bar{\boldsymbol{\omega}}(\tau)d\tau. \end{aligned} \tag{2.12}$$

The mean value and nuclear covariance of the initial state $\boldsymbol{x}_0$ are as follows:

$$\bar{\boldsymbol{x}}_0 = \begin{bmatrix} \bar{\boldsymbol{z}}(0) \\ \bar{\boldsymbol{z}} \end{bmatrix}, \qquad \boldsymbol{P}_0 = \mathcal{L}\mathcal{L}^*. \tag{2.13}$$

Remark 2.1. Note that the weighted scalar product (2.3), (2.4) assures that there exist real α such that $\boldsymbol{A}-\alpha\boldsymbol{I}$ has the nice property to be dissipative [3]. This property is used in the paper to prove the convergence of the approximation scheme.

For the reader's convenience, the definitions of some operators related to the system (2.5), (2.6) that will be extensively used in the paper are reported below.

PROPOSITION 2.2. *The operators* $\boldsymbol{B}^*$, $\boldsymbol{C}^*$, $\boldsymbol{F}^*$, $\boldsymbol{B}\boldsymbol{B}^*$, $\boldsymbol{C}^*\boldsymbol{C}$, $\boldsymbol{F}\boldsymbol{F}^*$, *and* $\boldsymbol{A}^*$ *are as follows:*

$$\boldsymbol{B}^* : \boldsymbol{M}_2 \mapsto \mathbb{R}^p, \qquad \boldsymbol{B}^* \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix} = \boldsymbol{B}_0^{\mathrm{T}} \boldsymbol{x}_0; \tag{2.14}$$

$$\boldsymbol{C}^* : \mathbb{R}^q \mapsto \boldsymbol{M}_2, \qquad \boldsymbol{C}^* \boldsymbol{y} = \begin{bmatrix} \boldsymbol{C}_0^{\mathrm{T}} \boldsymbol{y} \\ 0 \end{bmatrix}; \tag{2.15}$$

$$\boldsymbol{F}^* : \boldsymbol{M}_2 \mapsto \mathbb{R}^s, \qquad \boldsymbol{F}^* \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_0^{\mathrm{T}} \boldsymbol{x}_0 \\ \boldsymbol{0} \end{bmatrix}; \tag{2.16}$$

$$\boldsymbol{B}\boldsymbol{B}^* : \boldsymbol{M}_2 \mapsto \boldsymbol{M}_2, \qquad \boldsymbol{B}\boldsymbol{B}^* \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_0 \boldsymbol{B}_0^{\mathrm{T}} \boldsymbol{x}_0 \\ 0 \end{bmatrix}; \tag{2.17}$$

$$\boldsymbol{C}^*\boldsymbol{C} : \boldsymbol{M}_2 \mapsto \boldsymbol{M}_2, \qquad \boldsymbol{C}^*\boldsymbol{C} \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_0^{\mathrm{T}} \boldsymbol{C}_0 \boldsymbol{x}_0 \\ 0 \end{bmatrix}; \tag{2.18}$$

$$\boldsymbol{F}\boldsymbol{F}^* : \boldsymbol{M}_2 \mapsto \boldsymbol{M}_2, \qquad \boldsymbol{F}\boldsymbol{F}^* \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_0 \boldsymbol{F}_0^{\mathrm{T}} \boldsymbol{x}_0 \\ 0 \end{bmatrix}; \tag{2.19}$$

$$\begin{aligned} &\boldsymbol{A}^* : \mathcal{D}(\boldsymbol{A}^*) \mapsto \boldsymbol{M}_2, \\ &\boldsymbol{A}^* \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{bmatrix} = \begin{bmatrix} \delta\, \boldsymbol{y}_1(0) + \boldsymbol{A}_0^{\mathrm{T}} \boldsymbol{y}_0 \\ \dfrac{1}{g} \boldsymbol{A}_{01}^{\mathrm{T}} \boldsymbol{y}_0 - \dfrac{d}{d\vartheta} \left(\boldsymbol{y}_1 - \displaystyle\sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0, \boldsymbol{y}_1) \chi_{[-r,-r_j]} \right) \end{bmatrix}, \end{aligned} \tag{2.20}$$

with dense domain

$$\mathcal{D}(\boldsymbol{A}^*) = \left\{ \begin{bmatrix} \boldsymbol{y}_0 \\ \\ \boldsymbol{y}_1 \end{bmatrix} \middle| \begin{matrix} \boldsymbol{y}_0 \in \mathbb{R}^N, \quad \boldsymbol{A}_\delta^{\mathrm{T}} \boldsymbol{y}_0 = \boldsymbol{y}_1(-r), \\ \left(\boldsymbol{y}_1 - \displaystyle\sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0, \boldsymbol{y}_1) \chi_{[-r,-r_j]} \right) \in W^{1,2} \end{matrix} \right\}, \tag{2.21}$$

where

$$\boldsymbol{k}_j(\boldsymbol{y}_0, \boldsymbol{y}_1) = \frac{\boldsymbol{y}_1(-r_j) - \boldsymbol{A}_j^{\mathrm{T}} \boldsymbol{y}_0}{\delta - j + 1}, \qquad j = 1, \dots, \delta - 1. \tag{2.22}$$

The proof that the operator defined by (2.20), (2.21), (2.22) is in fact the adjoint of operator $\boldsymbol{A}$ is reported in appendix.

Remark 2.3. The difference between the case of just one pure delay and of multiple pure delays is given by summations in (2.20), (2.21), which vanish in the first case and complicate the analysis very much in the second one.

3. The finite horizon LQG for delay systems. In this section the problem of defining a feedback control law for the stochastic delay system (2.1), (2.2) is considered. In particular we are interested in the problem of synthesizing the control law that minimizes the cost functional

$$J_f(\boldsymbol{u}) = \int_0^{t_f} E[\boldsymbol{z}^{\mathrm{T}}(t)\boldsymbol{Q}_0\boldsymbol{z}(t) + \boldsymbol{u}^{\mathrm{T}}(t)\boldsymbol{u}(t)]dt, \tag{3.1}$$

with $0 < t_f < \infty$, where matrix $\boldsymbol{Q}_0$ is symmetric nonnegative definite. It can be readily recognized that the functional (3.1) admits the following representation in $\boldsymbol{M}_2$:

$$J_f(\boldsymbol{u}) = \int_0^{t_f} E[\big(\boldsymbol{Q}\boldsymbol{x}(t), \boldsymbol{x}(t)\big) + \boldsymbol{u}^{\mathrm{T}}(t)\boldsymbol{u}(t)]dt, \tag{3.2}$$

where $\boldsymbol{Q} : \boldsymbol{M}_2 \mapsto \boldsymbol{M}_2$ is defined as

$$\boldsymbol{Q}\begin{bmatrix}\boldsymbol{x}_0\\ \boldsymbol{x}_1\end{bmatrix} = \begin{bmatrix}\boldsymbol{Q}_0\boldsymbol{x}_0\\ \boldsymbol{0}\end{bmatrix} \tag{3.3}$$

and $\boldsymbol{x}(t)$ satisfies system equations (2.5), (2.6). The solution of this problem, as is well known, is the classical LQG controller given by the following equations [1]:

$$\boldsymbol{u}(t) = -\boldsymbol{B}^*\boldsymbol{R}(t_f - t)\hat{\boldsymbol{x}}(t), \tag{3.4}$$

$$\boldsymbol{R}(t) = \int_0^t \boldsymbol{T}^*(t-\tau)[\boldsymbol{Q} - \boldsymbol{R}(\tau)\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{R}(\tau)]\boldsymbol{T}(t-\tau)d\tau, \tag{3.5}$$

$$\hat{\boldsymbol{x}}(t) = \boldsymbol{T}(t)\hat{\boldsymbol{x}}_0 + \int_0^t \boldsymbol{T}(t-\tau)\left[\boldsymbol{P}(\tau)\boldsymbol{C}^*[\boldsymbol{y}(\tau) - \boldsymbol{C}\hat{\boldsymbol{x}}(\tau)] + \boldsymbol{B}\boldsymbol{u}(\tau)\right]d\tau, \tag{3.6}$$

$$\boldsymbol{P}(t) = \boldsymbol{T}(t)\boldsymbol{P}_0\boldsymbol{T}^*(t) + \int_0^t \boldsymbol{T}(t-\tau)[\boldsymbol{F}\boldsymbol{F}^* - \boldsymbol{P}(\tau)\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{P}(\tau)]\boldsymbol{T}^*(t-\tau)d\tau, \tag{3.7}$$

where $\boldsymbol{T}(t)$ is the semigroup governing the system, that is, the semigroup generated by the operator $\boldsymbol{A}$ in (2.7), (2.8), and $\hat{\boldsymbol{x}}_0$ and $\boldsymbol{P}_0$ are the expected value and the covariance operator of the initial state $\boldsymbol{x}(0)$ in $\boldsymbol{M}_2$, respectively. The solution given by these equations is a very important result only from a theoretical point of view. For our purposes we need to recall that the solutions of the Riccati equations (3.5), (3.7) evolve in the Hilbert space of Hilbert–Schmidt operators and moreover, for every $t_f < \infty$, there exist constants K_P and K_R such that [20]

$$\begin{aligned}\sup_{t\in[0,t_f]} \|\boldsymbol{P}(t)\|_{H.S.} = K_P < \infty,\\ \sup_{t\in[0,t_f]} \|\boldsymbol{R}(t)\|_{H.S.} = K_R < \infty,\end{aligned} \tag{3.8}$$

where, as usual, $\|\cdot\|_{H.S.}$ denotes the Hilbert–Schmidt norm [1].

In engineering applications, due to its infinite dimensional nature, such a solution is not directly implementable. Therefore it becomes important to investigate when such a solution admits a finite dimensional approximation.

Throughout the paper, given a Hilbert space $\mathcal{X}$ and a closed subspace $\mathcal{S} \subset \mathcal{X}$, the orthogonal projection operator from $\mathcal{X}$ to $\mathcal{S}$ will be denoted as $\boldsymbol{\Pi}_{\mathcal{S}}$.

In the next lemma, the linear space of bounded operators on a Hilbert space H is denoted $L(H)$.

LEMMA 3.1. *Let H_1, H_2 be separable Hilbert spaces. Let $\{G_m(t),\ t \in [0,t_f]\}$ be a sequence of strongly continuous $L(H_2)$ valued functions, strongly convergent to $\{G(t),\ t \in [0,t_f]\}$, uniformly on $[0,t_f]$. Let K be a compact subset in the Hilbert space of Hilbert–Schmidt operators mapping H_1 to H_2.*

Then $\|G_m(t)N - G(t)N\|_{H.S.}$ converges to zero, uniformly with respect to $N \in K$ and $t \in [0,t_f]$.

Proof. See [20]. □

LEMMA 3.2. *Let H_1 and H_2 be separable Hilbert spaces. Let $G(t)$ be a semigroup on H_2 and $G_n(t)$ a sequence of semigroups on H_2 strongly convergent to $G(t)$ uniformly with respect to $t \in [0,t_f]$. For $0 \le \tau \le t$, let $\Gamma(t,\tau)$ be the mild evolution operator*

$$\Gamma(t,\tau) = G(t-\tau) + \int_\tau^t G(t-\vartheta)Op(\vartheta)\Gamma(\vartheta,\tau)d\vartheta, \tag{3.9}$$

where $Op \in C([0,t_f];L(H_2))$ and let $\Gamma_n(t,\tau)$ be the sequence of mild evolution operators

$$\Gamma_n(t,\tau) = G_n(t-\tau) + \int_\tau^t G_n(t-\vartheta)Op_n(\vartheta)\Gamma_n(\vartheta,\tau)d\vartheta, \tag{3.10}$$

where $Op_n \in C([0,t_f];L(H_2))$ converges pointwise strongly to Op, uniformly in $[0,t_f]$. Let K be a compact subset in the Hilbert space of Hilbert–Schmidt operators mapping H_1 to H_2.

Then $\|\Gamma(t,\tau)N - \Gamma_n(t,\tau)N\|_{H.S.}$ converges to zero, uniformly with respect to $N \in K$ and $0 \le \tau \le t \le t_f$.

Proof. It is

$$\begin{aligned}
\|\Gamma(t,\tau)N - \Gamma_n(t,\tau)N\|_{H.S.} &\le \|G(t-\tau)N - G_n(t-\tau)N\|_{H.S.} \\
&+ \int_\tau^t \|G(t-\vartheta)Op(\vartheta)\| \cdot \|\Gamma(\vartheta,\tau)N - \Gamma_n(\vartheta,\tau)N\|_{H.S.}d\vartheta \\
&+ \int_\tau^t \|G(t-\vartheta)Op(\vartheta) - G_n(t-\vartheta)Op_n(\vartheta)\| \\
&\quad \cdot \|\Gamma_n(\vartheta,\tau)N - \Gamma(\vartheta,\tau)N\|_{H.S.}d\vartheta \\
&+ \int_\tau^t \|(G(t-\vartheta)Op(\vartheta) - G_n(t-\vartheta)Op_n(\vartheta))\Gamma(\vartheta,\tau)N\|_{H.S.}d\vartheta.
\end{aligned} \tag{3.11}$$

Let M be a positive real such that

$$\begin{aligned}
M &\ge \sup_{(t,\vartheta)\in[0,t_f]\times[0,t_f]} \|G(t)\|\|Op(\vartheta)\|, \\
M &\ge \sup_{(t,\vartheta,n)\in[0,t_f]\times[0,t_f]\times Z^+} \|G_n(t)\|\|Op_n(\vartheta)\|.
\end{aligned} \tag{3.12}$$

Then

$$\begin{aligned}\|\Gamma(t,\tau)N - \Gamma_n(t,\tau)N\|_{H.S.} \leq{}& \|G(t-\tau)N - G_n(t-\tau)N\|_{H.S.}\\ &+ \int_\tau^t \|(G(t-\vartheta)Op(\vartheta) - G_n(t-\vartheta)Op_n(\vartheta))\Gamma(\vartheta,\tau)N\|_{H.S.} d\vartheta\\ &+3M\int_\tau^t \|\big(\Gamma_n(\vartheta,\tau) - \Gamma(\vartheta,\tau)\big)N\|_{H.S.} d\vartheta.\end{aligned}\tag{3.13}$$

Applying the Gronwall's inequality,

$$\begin{aligned}\|\Gamma(t,\tau)N - \Gamma_n(t,\tau)N\|_{H.S.} \leq{}& e^{3Mt_f}\Big(\|G(t-\tau)N - G_n(t-\tau)N\|_{H.S.}\\ &+ \int_\tau^t \big\|\big(G(t-\vartheta)Op(\vartheta) - G_n(t-\vartheta)Op_n(\vartheta)\big)\Gamma(\vartheta,\tau)N\big\|_{H.S.} d\vartheta\Big)\\ &\leq e^{3Mt_f}\Big(\|G(t-\tau)N - G_n(t-\tau)N\|_{H.S.}\\ &+ \int_\tau^t \big\|\big(G(t-\vartheta) - G_n(t-\vartheta)\big)Op(\vartheta)\Gamma(\vartheta,\tau)N\big\|_{H.S.} d\vartheta\\ &+ \int_\tau^t M\big\|\big(Op(\vartheta) - Op_n(\vartheta)\big)\Gamma(\vartheta,\tau)N\big\|_{H.S.} d\vartheta\Big).\end{aligned}\tag{3.14}$$

Since the set of operators $\{Op(\vartheta)\Gamma(t,\tau)N, \vartheta \in [0,t_f], 0 \leq \tau \leq t \leq t_f\}$ and the set $\{\Gamma(t,\tau)N,\ 0 \leq \tau \leq t \leq t_f\}$ are compact in the Hilbert space of Hilbert–Schmidt operators mapping H_1 to H_2, by Lemma 3.1 the right-hand side of inequality (3.14) tends to zero for $n \to \infty$, and the lemma is proved. ∎

THEOREM 3.3. *Let Ψ_n and Ψ'_n be sequences of finite dimensional subspaces of $\boldsymbol{M}_2$ contained in $\mathcal{D}(\boldsymbol{A})$ and in $\mathcal{D}(\boldsymbol{A}^*)$, respectively. Let $\boldsymbol{\Pi}_{\Psi_n} : \boldsymbol{M}_2 \mapsto \Psi_n$ and $\boldsymbol{\Pi}_{\Psi'_n} : \boldsymbol{M}_2 \mapsto \Psi'_n$ be the sequences of orthoprojection operators in Ψ_n and Ψ'_n, respectively. Let $\boldsymbol{T}_{\Psi_n}(t)$ be the semigroup generated by the operator $\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{A}\boldsymbol{\Pi}_{\Psi_n} : \boldsymbol{M}_2 \mapsto \Psi_n$ and $\boldsymbol{T}^*_{\Psi'_n}(t)$ the semigroup generated by the operator $\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{A}^*\boldsymbol{\Pi}_{\Psi'_n} : \boldsymbol{M}_2 \mapsto \Psi'_n$. Let $\boldsymbol{P}_n(t)$ and $\boldsymbol{R}_n(t)$ be the solutions of the finite dimensional differential Riccati equations*

$$\begin{aligned}\dot{\boldsymbol{P}}_n(t) &= \boldsymbol{\Pi}_{\Psi_n}\boldsymbol{A}\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}_n(t) + \boldsymbol{P}_n(t)(\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{A}\boldsymbol{\Pi}_{\Psi_n})^*\\ &\qquad -\boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}_n(t) + \boldsymbol{\Pi}_{\Psi_n}\boldsymbol{F}\boldsymbol{F}^*\boldsymbol{\Pi}_{\Psi_n},\\ \boldsymbol{P}_n(0) &= \boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}_0\boldsymbol{\Pi}_{\Psi_n},\end{aligned}\tag{3.15}$$

$$\begin{aligned}\dot{\boldsymbol{R}}_n(t) &= \boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{A}^*\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{R}_n(t) + \boldsymbol{R}_n(t)(\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{A}^*\boldsymbol{\Pi}_{\Psi'_n})^*\\ &\qquad -\boldsymbol{R}_n(t)\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{R}_n(t) + \boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{Q}\boldsymbol{\Pi}_{\Psi'_n},\\ \boldsymbol{R}_n(0) &= \boldsymbol{0}.\end{aligned}\tag{3.16}$$

Assume the following hypotheses are satisfied:

(Hp_1) $\boldsymbol{\Pi}_{\Psi_n}$ *converges strongly to the identity operator;*

(Hp_2) $\boldsymbol{\Pi}_{\Psi'_n}$ *converges strongly to the identity operator;*

(Hp_3) $\boldsymbol{T}_{\Psi_n}(t)$ *converges strongly to $\boldsymbol{T}(t)$ uniformly in $[0,t_f]$;*

(Hp_4) $\boldsymbol{T}^*_{\Psi'_n}(t)$ *converges strongly to $\boldsymbol{T}^*(t)$ uniformly in $[0,t_f]$.*

Then

$$\begin{aligned} &\|\boldsymbol{P}_n(t) - \boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}\|_{H.S.} \to 0, \\ &\|\boldsymbol{R}_n(t) - \boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{R}(t)\boldsymbol{\Pi}_{\Psi'_n}\|_{H.S.} \to 0, \end{aligned} \qquad \textit{uniformly in} \quad [0, t_f]. \tag{3.17}$$

Proof. See the proof of Theorem 3 in [20]. □

Remark 3.4. Note that, with the given definitions, in general the semigroup $\boldsymbol{T}^*_{\Psi'_n}(t)$ generated by the operator $\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{A}^*\boldsymbol{\Pi}_{\Psi'_n}$ is different from the semigroup $\boldsymbol{T}^*_{\Psi_n}(t)$, the adjoint of the semigroup generated by $\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{A}\boldsymbol{\Pi}_{\Psi_n}$.

LEMMA 3.5. *Let Ψ_n and Ψ'_n be sequences of finite dimensional subspaces of $\boldsymbol{M}_2$ contained in $\mathcal{D}(\boldsymbol{A})$ and in $\mathcal{D}(\boldsymbol{A}^*)$, respectively. Let $\boldsymbol{\Pi}_{\Psi_n} : \boldsymbol{M}_2 \mapsto \Psi_n$ and $\boldsymbol{\Pi}_{\Psi'_n} : \boldsymbol{M}_2 \mapsto \Psi'_n$ be the corresponding sequences of orthoprojection operators. Let $\mathcal{H} = \boldsymbol{M}_2 \times \boldsymbol{M}_2$ and $\mathcal{H}_n = \boldsymbol{M}_2 \times \Psi_n$. Consider the following operators:*

$$\mathcal{A} = \begin{bmatrix} \boldsymbol{A} & 0 \\ 0 & \boldsymbol{A} \end{bmatrix} : \mathcal{D}(\boldsymbol{A}) \times \mathcal{D}(\boldsymbol{A}) \mapsto \mathcal{H}, \tag{3.18}$$

$$\mathcal{A}_n = \begin{bmatrix} \boldsymbol{A} & 0 \\ 0 & \boldsymbol{\Pi}_{\Psi_n}\boldsymbol{A}\boldsymbol{\Pi}_{\Psi_n} \end{bmatrix} : \mathcal{D}(\boldsymbol{A}) \times \boldsymbol{M}_2 \mapsto \mathcal{H}_n, \tag{3.19}$$

$$\boldsymbol{D}(t) = \begin{bmatrix} 0 & -\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{R}(t_f - t) \\ \boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{C} & -\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{R}(t_f - t) - \boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{C} \end{bmatrix} : \mathcal{H} \mapsto \mathcal{H}, \tag{3.20}$$

$$\boldsymbol{D}_n(t) = \begin{bmatrix} 0 & -\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{R}_n(t_f - t)\boldsymbol{\Pi}_{\Psi'_n} \\ \boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C} & -\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{R}_n(t_f - t)\boldsymbol{\Pi}_{\Psi'_n} - \boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{\Pi}_{\Psi_n} \end{bmatrix} :$$

$$\mathcal{H} \mapsto \mathcal{H}_n, \tag{3.21}$$

$$\boldsymbol{O}(t) = \begin{bmatrix} \boldsymbol{F} \\ \boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{G} \end{bmatrix} : \mathbb{R}^s \mapsto \mathcal{H}, \tag{3.22}$$

$$\boldsymbol{O}_n(t) = \begin{bmatrix} \boldsymbol{F} \\ \boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{G} \end{bmatrix} : \mathbb{R}^s \mapsto \mathcal{H}_n. \tag{3.23}$$

Let $\boldsymbol{S}(t)$ and $\boldsymbol{S}_n(t)$ be the semigroups generated by operators $\mathcal{A}$ and $\mathcal{A}_n$, respectively.

Let $\boldsymbol{\Pi}_{\mathcal{H}_n}$ be the following sequence of orthoprojection operators, strongly converging to identity,

$$\boldsymbol{\Pi}_{\mathcal{H}_n} = \begin{bmatrix} \boldsymbol{I} & 0 \\ 0 & \boldsymbol{\Pi}_{\Psi_n} \end{bmatrix} : \mathcal{H} \mapsto \mathcal{H}_n \tag{3.24}$$

Assume that assumptions Hp_1–Hp_4 of Theorem 3.3 *are satisfied.*

Then

(Th$_1$) $\boldsymbol{S}_n(t)$ *converges strongly to* $\boldsymbol{S}(t)$ *uniformly in* $[0, t_f]$;

(Th$_2$) $\|\boldsymbol{\Pi}_{\mathcal{H}_n}\boldsymbol{O}(t) - \boldsymbol{O}_n(t)\|_{H.S.} \to 0$ *uniformly in* $[0, t_f]$;

(Th$_3$) $\|\boldsymbol{\Pi}_{\mathcal{H}_n}\boldsymbol{D}(t) - \boldsymbol{D}_n(t)\|_{H.S.} \to 0$ *uniformly in* $[0, t_f]$.

Proof. Thesis Th$_1$ is an immediate consequence of hypothesis Hp$_3$.

As far as Th_2 is concerned, we have

$$\begin{aligned}\|\boldsymbol{\Pi}_{\mathcal{H}_n}\boldsymbol{O}(t)-\boldsymbol{O}_n(t)\|_{H.S.} &= \|\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{G}-\boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{G}\|_{H.S.}\\ &\le \Big\|\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{G}-\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{G}\\ &\qquad +\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{G}-\boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{G}\|_{H.S.}\\ &\le \|\boldsymbol{\Pi}_{\Psi_n}\|\|\boldsymbol{P}(t)\|_{H.S.}\|\boldsymbol{C}^*\boldsymbol{G}-\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{G}\|_{H.S.}\\ &\qquad +\|\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}-\boldsymbol{P}_n(t)\|_{H.S.}\|\boldsymbol{\Pi}_{\Psi_n}\|\|\boldsymbol{C}^*\boldsymbol{G}\|.\end{aligned}\tag{3.25}$$

From (3.25) it follows that $\|\boldsymbol{\Pi}_{\mathcal{H}_n}\boldsymbol{O}(t)-\boldsymbol{O}_n(t)\|_{H.S.}\to 0$ uniformly with respect to $t\in[0,t_f]$ because of the boundedness of $\|\boldsymbol{P}(t)\|_{H.S}$ and the uniform convergence of $\boldsymbol{P}_n(t)$ stated in Theorem 3.3.

As for thesis Th_3, it is

$$\boldsymbol{\Pi}_{\mathcal{H}_n}\boldsymbol{D}(t)-\boldsymbol{D}_n(t)=\begin{bmatrix}0 & \boldsymbol{O}p^n_{1,2}(t)\\ \boldsymbol{O}p^n_{2,1}(t) & \boldsymbol{O}p^n_{2,2}(t)\end{bmatrix},\tag{3.26}$$

where

$$\begin{aligned}&\boldsymbol{O}p^n_{1,2}(t)=\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{R}_n(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}-\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{R}(t_f-t);\\ &\boldsymbol{O}p^n_{2,1}(t)=\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{C}-\boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C};\\ &\boldsymbol{O}p^n_{2,2}(t)=\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{R}_n(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}\\ &\qquad +\boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{\Pi}_{\Psi_n}-\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{R}(t_f-t)-\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{C}.\end{aligned}\tag{3.27}$$

To prove Th_3 it is sufficient to prove that the three operators in (3.27) converge uniformly to zero in the $H.S.$ norm. Let us start with operator $\boldsymbol{O}p^n_{1,2}(t)$. We have

$$\begin{aligned}\|\boldsymbol{O}p^n_{1,2}(t)\| &\le \|\boldsymbol{B}\boldsymbol{B}^*\|\big\|\boldsymbol{R}_n(t_f-t)-\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{R}(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}\big\|_{H.S.}\|\boldsymbol{\Pi}_{\Psi'_n}\|\\ &\qquad +\|\boldsymbol{B}\boldsymbol{B}^*\|\big\|\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{R}(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}-\boldsymbol{R}(t_f-t)\big\|_{H.S.}.\end{aligned}\tag{3.28}$$

Moreover, from the uniform convergence of $\boldsymbol{R}_n(t)$ stated in Theorem 3.3, $\boldsymbol{R}$ being self-adjoint, and for Lemma 3.1, by

$$\begin{aligned}&\|\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{R}(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}-\boldsymbol{R}(t_f-t)\|_{H.S.}\\ &\quad\le \|\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{R}(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}-\boldsymbol{R}(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}+\boldsymbol{R}(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}-\boldsymbol{R}(t_f-t)\|_{H.S.}\\ &\quad\le 2\|\boldsymbol{\Pi}_{\Psi'_n}\boldsymbol{R}(t_f-t)-\boldsymbol{R}(t_f-t)\|,\end{aligned}\tag{3.29}$$

it follows that $\|\boldsymbol{O}p^n_{1,2}(t)\|_{H.S.}\to 0$ uniformly in $[0,t_f]$. Consider now the term $\boldsymbol{O}p^n_{2,1}(t)$. Its Hilbert–Schmidt norm satisfies

$$\begin{aligned}&\|\boldsymbol{O}p^n_{2,1}(t)\|_{H.S.}\\ &\quad\le \big\|\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)-\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}+\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}-\boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\big\|_{H.S.}\|\boldsymbol{C}^*\boldsymbol{C}\|\\ &\quad\le \Big(\|\boldsymbol{P}(t)-\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}\|_{H.S.}+\|(\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}-\boldsymbol{P}_n(t))\boldsymbol{\Pi}_{\Psi_n}\|_{H.S.}\Big)\|\boldsymbol{C}^*\boldsymbol{C}\|.\end{aligned}\tag{3.30}$$

Since, by Lemma 3.1, $\|\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)-\boldsymbol{P}(t)\|_{H.S.}\to 0$ uniformly and $\|(\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}-\boldsymbol{P}_n(t))\boldsymbol{\Pi}_{\Psi_n}\|_{H.S.}\to 0$ by Theorem 3.3, it follows that the norm of $\boldsymbol{O}p_{2,1}^n(t)$ tends to zero uniformly in $[0,t_f]$.

It remains to prove that $\|\boldsymbol{O}p_{2,2}^n(t)\|_{H.S.}\to 0$ uniformly.

$$\begin{aligned}\|\boldsymbol{O}p_{2,2}^n(t)\|_{H.S.} &\le \|\boldsymbol{\Pi}_{\Psi_n}\|\big\|\boldsymbol{BB}^*\boldsymbol{R}(t_f-t)-\boldsymbol{BB}^*\boldsymbol{R}_n(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}\big\|_{H.S.}\\ &\quad+\big\|\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{C}-\boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{\Pi}_{\Psi_n}\big\|_{H.S.}\end{aligned} \tag{3.31}$$

Uniform convergence to zero of $\|\boldsymbol{BB}^*\boldsymbol{R}(t_f-t)-\boldsymbol{BB}^*\boldsymbol{R}_n(t_f-t)\boldsymbol{\Pi}_{\Psi'_n}\|_{H.S.}$ has already been proved. Moreover,

$$\begin{aligned}&\|\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{C}-\boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{\Pi}_{\Psi_n}\|_{H.S.}\\ &\quad\le \|\boldsymbol{\Pi}_{\Psi_n}\|\|\boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{C}-\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{\Pi}_{\Psi_n}\|_{H.S.}\\ &\qquad+\big\|\big(\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}-\boldsymbol{P}_n(t)\big)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C}\big\|_{H.S.}\|\boldsymbol{\Pi}_{\Psi_n}\|.\end{aligned} \tag{3.32}$$

Again, as proved in [20], the term $\|(\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{\Pi}_{\Psi_n}-\boldsymbol{P}_n(t))\boldsymbol{\Pi}_{\Psi_n}\|_{H.S.}\to 0$ uniformly and thanks to Lemma 3.1 also $\|\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{P}(t)-\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\|_{H.S.}\to 0$ uniformly. From (3.32) it follows that $\|\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{P}(t)\boldsymbol{C}^*\boldsymbol{C}-\boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{\Pi}_{\Psi_n}\|_{H.S.}\to 0$ uniformly, so that $\|\boldsymbol{O}p_{2,2}^n(t)\|_{H.S.}\to 0$ uniformly in $[0,t_f]$, and the lemma is proved. ☐

LEMMA 3.6. *Let $\boldsymbol{U}(t,\tau), 0\le\tau\le t$, be the mild evolution operator*

$$\boldsymbol{U}(t,\tau)=\boldsymbol{S}(t-\tau)+\int_\tau^t \boldsymbol{S}(t-\vartheta)\boldsymbol{D}(\vartheta)\boldsymbol{U}(\vartheta,\tau)d\vartheta. \tag{3.33}$$

Let $\{\boldsymbol{U}_n(t,\tau)\}$ be the sequence of mild evolution operators

$$\boldsymbol{U}_n(t,\tau)=\boldsymbol{S}_n(t-\tau)+\int_\tau^t \boldsymbol{S}_n(t-\vartheta)\boldsymbol{D}_n(\vartheta)\boldsymbol{U}_n(\vartheta,\tau)d\vartheta. \tag{3.34}$$

Then $\{\boldsymbol{U}_n(t,\tau)\}$ converges strongly to $\boldsymbol{U}(t,\tau)$ uniformly in $0\le\tau\le t\le t_f$, that is, given any $X\in\mathcal{H}$,

$$\lim_{n\to\infty}\sup_{0\le\tau\le t\le t_f}\|\boldsymbol{U}_n(t,\tau)X-\boldsymbol{U}(t,\tau)X\|=0. \tag{3.35}$$

Proof. Let us denote by $g(t,\tau)$ and $g_n(t,\tau)$ the quantities

$$g(t,\tau)=\boldsymbol{U}(t,\tau)X, \tag{3.36}$$

$$g_n(t,\tau)=\boldsymbol{U}_n(t,\tau)X, \tag{3.37}$$

from which, denoting the approximation error by $e_n(t,\tau)$

$$e_n(t,\tau)=g(t,\tau)-g_n(t,\tau), \tag{3.38}$$

we have

$$\begin{aligned}e_n(t,\tau)&=\boldsymbol{S}(t-\tau)X+\int_\tau^t \boldsymbol{S}(t-\vartheta)\boldsymbol{D}(\vartheta)g(\vartheta,\tau)d\vartheta\\ &\quad-\boldsymbol{S}_n(t-\tau)X-\int_\tau^t \boldsymbol{S}_n(t-\vartheta)\boldsymbol{D}_n(\vartheta)g_n(\vartheta,\tau)d\vartheta,\end{aligned} \tag{3.39}$$

and therefore

$$\begin{aligned} e_n(t,\tau) &= (\boldsymbol{S}(t-\tau) - \boldsymbol{S}_n(t-\tau))X \\ &\quad + \int_\tau^t \Big(\boldsymbol{S}(t-\vartheta)\boldsymbol{D}(\vartheta)g(\vartheta,\tau) - \boldsymbol{S}_n(t-\vartheta)\boldsymbol{D}_n(\vartheta)g_n(\vartheta,\tau)\Big)d\vartheta \end{aligned} \tag{3.40}$$

from which

$$\begin{aligned} \|e_n(t,\tau)\| &\le \|(\boldsymbol{S}(t-\tau) - \boldsymbol{S}_n(t-\tau))X\| \\ &\quad + \int_\tau^t \|\boldsymbol{S}(t-\vartheta)\boldsymbol{D}(\vartheta) - \boldsymbol{S}_n(t-\vartheta)\boldsymbol{D}(\vartheta)\|\|g(\vartheta,\tau)\|d\vartheta \\ &\quad + \int_\tau^t \|\boldsymbol{S}_n(t-\vartheta)\|\|\boldsymbol{\Pi}_{\mathcal{H}_n}\boldsymbol{D}(\vartheta) - \boldsymbol{D}_n(\vartheta)\|\|g(\vartheta,\tau)\|d\vartheta \\ &\quad + \int_\tau^t \|\boldsymbol{S}_n(t-\vartheta)\boldsymbol{D}_n(\vartheta)\|\|e_n(\vartheta,\tau)\|d\vartheta. \end{aligned} \tag{3.41}$$

Now, given $\epsilon > 0$, by Lemma 3.1 there exists an integer $\nu_{\epsilon,X}$ such that, for all $n > \nu_{\epsilon,X}$, we have

$$\|e_n(t,\tau)\| \le \epsilon + \bar{S}\bar{D}\int_\tau^t \|e_n(\vartheta,\tau)\|d\vartheta, \tag{3.42}$$

where

$$\begin{aligned} \bar{S} &= \sup_{n,t\in[0,t_f]} \|\boldsymbol{S}_n(t)\|, \\ \bar{D} &= \sup_{n,t\in[0,t_f]} \|\boldsymbol{D}_n(t)\|. \end{aligned} \tag{3.43}$$

By Gronwall's lemma,

$$\|e_n(t,\tau)\| \le \epsilon e^{\bar{S}\bar{D}(t-\tau)}, \tag{3.44}$$

and therefore

$$\sup_{0\le\tau\le t\le t_f} \|e_n(t,\tau)\| \le \epsilon e^{\bar{S}\bar{D}t_f}. \tag{3.45}$$

This concludes the proof. □

Now, the main theorem can be given.

THEOREM 3.7. *Using the same hypotheses of Theorem* 3.3, *let* $\boldsymbol{u}_n(t)$ *be the input obtained by the following finite dimensional equations:*

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}}_n(t) &= \boldsymbol{\Pi}_{\Psi_n}\boldsymbol{A}\boldsymbol{\Pi}_{\Psi_n}\hat{\boldsymbol{x}}_n(t) + \boldsymbol{\Pi}_{\Psi_n}\boldsymbol{B}\boldsymbol{u}_n(t) + \boldsymbol{P}_n(t)\boldsymbol{\Pi}_{\Psi_n}\boldsymbol{C}^*\big(\boldsymbol{y}(t) - \boldsymbol{C}\boldsymbol{\Pi}_{\Psi_n}\hat{\boldsymbol{x}}_n(t)\big), \\ \hat{\boldsymbol{x}}_n(0) &= \boldsymbol{\Pi}_{\Psi_n}\hat{\boldsymbol{x}}(0), \end{aligned} \tag{3.46}$$

$$\boldsymbol{u}_n(t) = -\boldsymbol{B}^*\boldsymbol{R}_n(t_f - t)\boldsymbol{\Pi}_{\Psi'_n}\hat{\boldsymbol{x}}_n(t), \tag{3.47}$$

where $\boldsymbol{P}_n$ *and* $\boldsymbol{R}_n$ *are given by* (3.15) *and* (3.16). *Let* $\boldsymbol{u}(t)$ *be the optimal input,* $\hat{\boldsymbol{x}}(t)$ *the optimal estimated state,* $\boldsymbol{x}_n(t)$ *and* $\boldsymbol{x}(t)$ *the actual state evolving when* $\boldsymbol{u}_n(t)$ *and* $\boldsymbol{u}(t)$ *are applied to system* (2.1), (2.2), *respectively.*

Then

$$\lim_{n\to\infty} E\|\boldsymbol{x}_n - \boldsymbol{x}\|^2_{L_2([0,t_f];\boldsymbol{M}_2)} = 0, \tag{3.48}$$

$$\lim_{n\to\infty} E\|\hat{\boldsymbol{x}}_n - \hat{\boldsymbol{x}}\|^2_{L_2([0,t_f];\boldsymbol{M}_2)} = 0, \tag{3.49}$$

$$\lim_{n\to\infty} E\|\boldsymbol{u}_n - \boldsymbol{u}\|^2_{L_2([0,t_f];\mathbb{R}^p)} = 0, \tag{3.50}$$

$$\lim_{n\to\infty} |J_f(\boldsymbol{u}_n) - J_f(\boldsymbol{u})| = 0. \tag{3.51}$$

Proof. Let $\boldsymbol{X} = \begin{bmatrix}\boldsymbol{x}\\ \hat{\boldsymbol{x}}\end{bmatrix}$ and $\boldsymbol{X}_n = \begin{bmatrix}\boldsymbol{x}_n\\ \hat{\boldsymbol{x}}_n\end{bmatrix}$. It is

$$\begin{aligned} \dot{\boldsymbol{X}}(t) &= \mathcal{A}\boldsymbol{X}(t) + \boldsymbol{D}(t)\boldsymbol{X}(t) + \boldsymbol{O}(t)\boldsymbol{\omega}(t),\\ \dot{\boldsymbol{X}}_n(t) &= \mathcal{A}_n\boldsymbol{X}_n(t) + \boldsymbol{D}_n(t)\boldsymbol{X}_n(t) + \boldsymbol{O}_n(t)\boldsymbol{\omega}(t),\\ \boldsymbol{X}(0) &= \begin{bmatrix}\boldsymbol{x}(0)\\ \hat{\boldsymbol{x}}(0)\end{bmatrix}, \qquad \boldsymbol{X}_n(0) = \begin{bmatrix}\boldsymbol{x}(0)\\ \boldsymbol{\Pi}_{\Psi_n}\hat{\boldsymbol{x}}(0)\end{bmatrix}, \end{aligned} \tag{3.52}$$

where $\mathcal{A}, \mathcal{A}_n, \boldsymbol{D}(t), \boldsymbol{D}_n(t), \boldsymbol{O}(t), \boldsymbol{O}_n(t)$ have been defined in Lemma 3.5.

Let $\boldsymbol{S}(t)$, $\boldsymbol{S}_n(t)$, $\boldsymbol{U}(t,\tau)$, $\boldsymbol{U}_n(t,\tau)$ be as in Lemmas 3.5, 3.6. We have

$$\boldsymbol{X}(t) = \boldsymbol{S}(t)\boldsymbol{X}(0) + \int_0^t \boldsymbol{S}(t-\tau)\big(\boldsymbol{D}(\tau)\boldsymbol{X}(\tau) + \boldsymbol{O}(\tau)\boldsymbol{\omega}(\tau)\big)d\tau, \tag{3.53}$$

$$\boldsymbol{X}_n(t) = \boldsymbol{S}_n(t)\boldsymbol{X}_n(0) + \int_0^t \boldsymbol{S}_n(t-\tau)\big(\boldsymbol{D}_n(\tau)\boldsymbol{X}_n(\tau) + \boldsymbol{O}_n(\tau)\boldsymbol{\omega}(\tau)\big)d\tau, \tag{3.54}$$

which can be rewritten as

$$\boldsymbol{X}(t) = \boldsymbol{U}(t,0)\boldsymbol{X}(0) + \int_0^t \boldsymbol{U}(t,\tau)\boldsymbol{O}(\tau)\boldsymbol{\omega}(\tau)d\tau, \tag{3.55}$$

$$\boldsymbol{X}_n(t) = \boldsymbol{U}_n(t,0)\boldsymbol{X}_n(0) + \int_0^t \boldsymbol{U}_n(t,\tau)\boldsymbol{O}_n(\tau)\boldsymbol{\omega}(\tau)d\tau. \tag{3.56}$$

Let us introduce the Hilbert spaces

$$\boldsymbol{W}_{\boldsymbol{X},t} = L_2([0,t];\mathcal{H}), \qquad \boldsymbol{W}_{\omega,t} = L_2([0,t];\mathbb{R}^s), \tag{3.57}$$

and define the operators

$$\begin{aligned} &\boldsymbol{L}_t : \boldsymbol{W}_{\omega,t} \mapsto \boldsymbol{W}_{\boldsymbol{X},t},\\ &f = \boldsymbol{L}_t\, g, \qquad f(\tau) = \int_0^\tau \boldsymbol{U}(\tau,\vartheta)\boldsymbol{O}(\vartheta)g(\vartheta)d\vartheta, \end{aligned} \tag{3.58}$$

$$\begin{aligned} &\boldsymbol{L}_{t,n} : \boldsymbol{W}_{\omega,t} \mapsto \boldsymbol{W}_{\boldsymbol{X},t},\\ &f = \boldsymbol{L}_{t,n}\, g, \qquad f(\tau) = \int_0^\tau \boldsymbol{U}_n(\tau,\vartheta)\boldsymbol{O}_n(\vartheta)g(\vartheta)d\vartheta, \end{aligned} \tag{3.59}$$

and the functions

$$\boldsymbol{X}_0 : \quad \boldsymbol{X}_0(\tau) = \boldsymbol{U}(\tau,0)\boldsymbol{X}(0), \tag{3.60}$$

$$\boldsymbol{X}_{0,n} : \quad \boldsymbol{X}_{0,n}(\tau) = \boldsymbol{U}_n(\tau,0)\boldsymbol{X}_n(0). \tag{3.61}$$

In the space $\boldsymbol{W}_{\boldsymbol{X},t}$, (3.55), (3.56) can be expressed as

$$\boldsymbol{X} = \boldsymbol{X}_0 + \boldsymbol{L}_t\boldsymbol{\omega}, \tag{3.62}$$

$$\boldsymbol{X}_n = \boldsymbol{X}_{0,n} + \boldsymbol{L}_{t,n}\boldsymbol{\omega}, \tag{3.63}$$

so that

$$\begin{aligned} E\|\boldsymbol{X} - \boldsymbol{X}_n\|^2_{\boldsymbol{W}_{\boldsymbol{X},t}} &= E\|\boldsymbol{X}_0 - \boldsymbol{X}_{0,n} + (\boldsymbol{L}_t - \boldsymbol{L}_{t,n})\boldsymbol{\omega}\|^2_{\boldsymbol{W}_{\boldsymbol{X},t}} \\ &\leq 2E\|\boldsymbol{X}_0 - \boldsymbol{X}_{0,n}\|^2_{\boldsymbol{W}_{\boldsymbol{X},t}} + 2E\|(\boldsymbol{L}_t - \boldsymbol{L}_{t,n})\boldsymbol{\omega}\|^2_{\boldsymbol{W}_{\boldsymbol{X},t}}. \end{aligned} \tag{3.64}$$

The first term in the right-hand side goes to zero thanks to Lemma 3.6.

For the second term we have

$$\begin{aligned} &E\|(\boldsymbol{L}_t - \boldsymbol{L}_{t,n})\boldsymbol{\omega}\|^2_{\boldsymbol{W}_{(\boldsymbol{X},t)}} = \|\boldsymbol{L}_t - \boldsymbol{L}_{t,n}\|^2_{H.S.} \\ &\quad = \int_0^t \int_0^\tau \left\|\boldsymbol{U}(\tau,\vartheta)\boldsymbol{O}(\vartheta) - \boldsymbol{U}_n(\tau,\vartheta)\boldsymbol{O}_n(\vartheta)\right\|^2_{H.S.} d\vartheta d\tau \\ &\quad \leq 2\int_0^t \int_0^\tau \left\|\big(\boldsymbol{U}(\tau,\vartheta) - \boldsymbol{U}_n(\tau,\vartheta)\big)\boldsymbol{O}(\vartheta)\right\|^2_{H.S.} d\vartheta d\tau \\ &\qquad + 2\int_0^t \int_0^\tau \left\|\boldsymbol{U}_n(\tau,\vartheta)\big(\boldsymbol{O}(\vartheta) - \boldsymbol{O}_n(\vartheta)\big)\right\|^2_{H.S.} d\vartheta d\tau \\ &\quad \leq \sup_{0\leq\vartheta\leq\tau\leq t} \ \sup_{\boldsymbol{M}\in\{\boldsymbol{O}(\vartheta),\vartheta\in[0,t]\}} \left\|\big(\boldsymbol{U}(\tau,\vartheta) - \boldsymbol{U}_n(\tau,\vartheta)\big)M\right\|^2_{H.S.} t^2 \\ &\qquad + \sup_{0\leq\vartheta\leq\tau\leq t,\ \ n\in Z^+} \|\boldsymbol{U}_n(\tau,\vartheta)\|t^2 \sup_{\vartheta\in[0,t]} \|\boldsymbol{O}(\vartheta) - \boldsymbol{O}_n(\vartheta)\|^2_{H.S.} \end{aligned} \tag{3.65}$$

which goes to zero by using Lemmas 3.1, 3.5, and 3.6. This concludes the proof. ∎

4. The approximation scheme. In this section, we will derive the approximation scheme for the LQG controller (3.4)–(3.7). The first step is the definition of the sequences Ψ_n and Ψ'_n of subspaces approximating $\mathcal{D}(\boldsymbol{A})$ and $\mathcal{D}(\boldsymbol{A}^*)$. This is made by a suitable definition of basis vectors for subspaces $\Psi_n \subset \mathcal{D}(\boldsymbol{A})$ and $\Psi'_n \subset \mathcal{D}(\boldsymbol{A}^*)$. In order to avoid confusion with the general settings in section 3, the forthcoming choice for Ψ_n and Ψ'_n will be denoted by Φ_n and Φ'_n respectively. In [3] the dynamics of linear delay systems is approximated using classical first order splines uniformly distributed over the interval $[-r, 0]$. With this choice the computation of matrix representation of the approximated operators is quite complex due to the fact that in general, for a given number n of subintervals of $[-r, 0]$, the delay instants $-r_j$ do not coincide with knots of splines.

It is useful to define a multi-index $s = (n_1, \ldots, n_\delta)$ that characterizes the partition of each interval $[-r_i, -r_{i-1}]$, for $i = 1, \ldots, \delta$, into n_i subintervals of length $(r_i - r_{i-1})/n_i$, in which $n_i + 1$ classical first order splines are considered (see Figure 1), numbered from 0 to n_i.

Definition 4.1. *A sequence $\{s_n\}$ of multi-indexes, defined for $n = 1, 2, \ldots$, where n is the lowest of indexes n_j of the multi-index (i.e., $n = \min\{s_n\}$), is denoted a test sequence if there exists a constant $\bar{c}$ such that for each n it is $\max\{s_n\}/n \leq \bar{c}$.*

Let $t^i_j = -r_{i-1} - (r_i - r_{i-1})j/n_i$, for $j = 0, 1, \ldots, n_i$, $i = 1, \ldots, \delta$. Let $spline^i_j$ be the spline j of interval i, that is, the spline with knot in t^i_j. Let ϕ_k, $k = 1, 2, \ldots, N$, be the canonical base in $\mathbb{R}^N$.

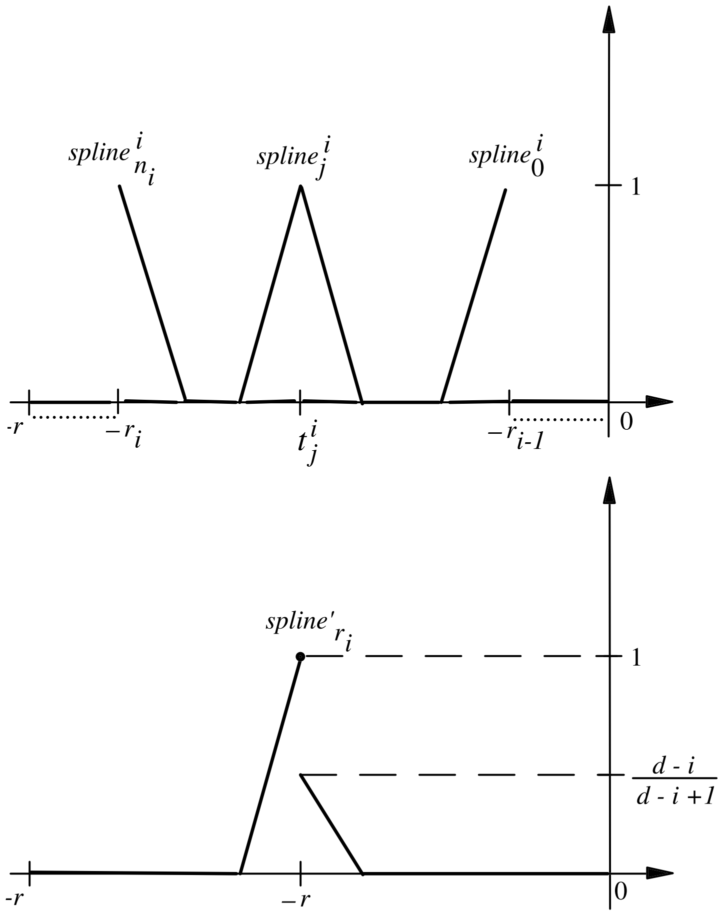

FIG. 1. *First order splines used for the approximation schemes.*

The approximating subspace Φ_n and Φ'_n are defined as follows.

DEFINITION 4.2. *For any given multi-index s_n of a test sequence let Φ_n be the subspace of linear combinations of vectors v_h^i, $v_k^{r_i}$ defined as follows:*

$$v_k^1 = \begin{bmatrix} \phi_k \\ \phi_k spline_0^1 \end{bmatrix}, \quad k = 1, \ldots, N, \tag{4.1}$$

$$v_{jN+k}^i = \begin{bmatrix} 0 \\ \phi_k spline_j^i \end{bmatrix}, \quad k = 1, \ldots, N, \quad j = 1, \ldots, n_i - 1, \tag{4.2}$$

$$v_{n_\delta N+k}^\delta = \begin{bmatrix} 0 \\ \phi_k spline_{n_\delta}^\delta \end{bmatrix}, \quad k = 1, \ldots, N, \tag{4.3}$$

$$v_k^{r_i} = \begin{bmatrix} 0 \\ \phi_k spline_{r_i} \end{bmatrix}, \quad \begin{matrix} i = 1, \ldots, \delta - 1, \\ k = 1, \ldots, N, \end{matrix} \tag{4.4}$$

where

$$spline_{r_i} = spline^i_{n_i} \cdot \chi_{(-r_i,-r_{i-1}]} + spline^{i+1}_0, \quad i = 1, \dots, \delta - 1. \tag{4.5}$$

DEFINITION 4.3. *For any given multi-index s_n of a test sequence let Φ'_n be the subspace of linear combinations of vectors w^i_h, $w^{r_i}_k$, w'_k defined as follows:*

$$w^1_k = \begin{bmatrix} 0 \\ \boldsymbol{\phi}_k spline^1_0 \end{bmatrix}, \qquad k = 1, \dots, N, \tag{4.6}$$

$$w^i_{jN+k} = v^i_{jN+k}, \qquad k = 1, \dots, N, \qquad \begin{matrix} i = 1, \dots, \delta, \\ j = 1, \dots, n_i - 1, \end{matrix} \tag{4.7}$$

$$w^{r_j}_k = \begin{bmatrix} 0 \\ \boldsymbol{\phi}_k spline'_{r_j} \end{bmatrix}, \quad j = 1, \dots, \delta - 1, \quad k = 1, \dots, N, \tag{4.8}$$

where

$$spline'_{r_j} = \frac{(\delta - j)}{(\delta - j + 1)} spline^j_{n_j} \cdot \chi_{(-r_j,-r_{j-1}]} + spline^{j+1}_0, \tag{4.9}$$

$$spline''_{r_j} = spline^j_{n_j} \cdot \chi_{(-r_j,-r_{j-1}]}, \qquad j = 1, \dots, \delta - 1, \tag{4.10}$$

$$w'_k = \begin{bmatrix} \phi_k \\ a_{\delta k} spline^\delta_{n_\delta} + \sum_{j=1}^{\delta-1} \frac{1}{(\delta-j+1)} a_{jk} spline''_{r_j} \end{bmatrix}, \quad k = 1, \dots, N, \tag{4.11}$$

where a_{jk} is the k column of matrix $\boldsymbol{A}^{\mathrm{T}}_j$ for $j = 1, \dots, \delta$.

THEOREM 4.4. *For each multi-index s_n of a test sequence, it is $\Phi_n \subset \mathcal{D}(\boldsymbol{A})$ and $\Phi'_n \subset \mathcal{D}(\boldsymbol{A}^*)$.*

Proof. It is immediate to verify that $\Phi_n \subset \mathcal{D}(\boldsymbol{A})$. A more detailed proof is required to show that $\Phi'_n \subset \mathcal{D}(\boldsymbol{A}^*)$. To this aim it is sufficient to verify that each vector w belongs to it. It is easy to check that vectors w^i_{jN+k}, for $i = 1, \dots, \delta$, $k = 1, \dots, N$, $j = 1, \dots, n_i - 1$, and vectors w^1_k, for $k = 1, \dots, N$, belong to $\boldsymbol{D}(\boldsymbol{A}^*)$. Let us consider now the vectors $w^{r_i}_k$, for $i = 1, \dots, \delta - 1$, $k = 1, \dots, N$ (as usual, we shall indicate the part in $\mathbb{R}^N$ by using the subscript 0 and the part in L_2 by using the subscript 1).

From the definition, for each k it is

$$\begin{aligned} (w^{r_i}_k)_0 &= 0, \\ (w^{r_i}_k)_1 (-r_j) &= 0, \quad i, j = 1, \dots, \delta - 1, \ i \neq j, \\ (w^{r_i}_k)_1 (-r) &= 0, \quad i = 1, \dots, \delta - 1, \end{aligned} \tag{4.12}$$

so that for $k = 1, \dots, N, \ \ i = 1, \dots, \delta - 1$,

$$(w^{r_i}_k)_1 (-r) = \boldsymbol{A}^{\mathrm{T}}_\delta (w^{r_i}_k)_0 \tag{4.13}$$

and

$$\begin{aligned} \sum_{j=1}^{\delta-1} \boldsymbol{k}_j ((w^{r_i}_k)_0, (w^{r_i}_k)_1) \chi_{[-r,-r_j]} &= \sum_{j=1}^{\delta-1} \frac{(w^{r_i}_k)_1 (-r_j) - \boldsymbol{A}^{\mathrm{T}}_j (w^{r_i}_k)_0}{\delta - j + 1} \chi_{[-r,-r_j]} \\ &= \frac{(w^{r_i}_k)_1 (-r_i)}{\delta - i + 1} \chi_{[-r,-r_i]}. \end{aligned} \tag{4.14}$$

So it is only to be verified that for $i = 1, \ldots, \delta - 1,\ k = 1, \ldots, N$,

$$(w_k^{r_i})_1 - \frac{1}{(\delta - i + 1)} (w_k^{r_i})_1 (-r_i) \chi_{[-r,-r_i]} \in W^{1,2}. \tag{4.15}$$

Since

$$spline'_{r_i}(-r_i) - \frac{1}{\delta - i + 1} = \frac{\delta - i}{\delta - i + 1} = \lim_{\vartheta \to -r_i^+} spline'_{r_i}(\vartheta), \tag{4.16}$$

(4.15) is clearly true.

For vectors w'_k defined in (4.11), for $k = 1, \ldots, N$, it is

$$(w'_k)_1(-r) = a_{\delta k} = \boldsymbol{A}_\delta^{\mathrm{T}} \phi_k = \boldsymbol{A}_\delta^{\mathrm{T}} (w'_k)_0. \tag{4.17}$$

It is also

$$(w'_k)_1(-r_i) = 0, \quad i = 1, \ldots, \delta - 1, \tag{4.18}$$

and therefore

$$\begin{aligned} \sum_{j=1}^{\delta-1} \boldsymbol{k}_j \left((w'_k)_0, (w'_k)_1\right) \chi_{[-r,-r_j]} &= \sum_{j=1}^{\delta-1} \frac{(w'_k)_1 (-r_j) - \boldsymbol{A}_j^{\mathrm{T}} (w'_k)_0}{\delta - j + 1} \chi_{[-r,-r_j]} \\ &= \sum_{j=1}^{\delta-1} \frac{-\boldsymbol{A}_j^{\mathrm{T}} (w'_k)_0}{(\delta - j + 1)} \chi_{[-r,-r_j]} = \sum_{j=1}^{\delta-1} \frac{-a_{jk}}{(\delta - j + 1)} \chi_{[-r,-r_j]}. \end{aligned} \tag{4.19}$$

From

$$\lim_{\vartheta \mapsto -r_j^+} (w'_k)_1 (\vartheta) = \frac{a_{jk}}{\delta - j + 1} \tag{4.20}$$

it follows, for $i = 1, \ldots, \delta - 1$,

$$\begin{aligned} &\left((w'_k)_1 + \sum_{j=1}^{\delta-1} \frac{a_{jk}}{(\delta - j + 1)} \chi_{[-r,-r_j]}\right)(-r_i) = (w'_k)_1(-r_i) + \sum_{j=1}^{i} \frac{a_{jk}}{(\delta - j + 1)} \\ &\quad = \sum_{j=1}^{i} \frac{a_{jk}}{\delta - j + 1} = \frac{a_{ik}}{\delta - i + 1} + \sum_{j=1}^{i-1} \frac{a_{jk}}{\delta - j + 1} \\ &\quad = \lim_{\vartheta \mapsto -r_i^+} (w'_k)_1 (\vartheta) + \lim_{\vartheta \mapsto -r_i^+} \sum_{j=1}^{\delta-1} \frac{a_{jk}}{\delta - j + 1} \chi_{[-r,-r_j]}(\vartheta), \end{aligned} \tag{4.21}$$

and so

$$\left((w'_k)_1 + \sum_{j=1}^{\delta-1} \frac{a_{jk}}{\delta - j + 1} \chi_{[-r,-r_j]}\right) \in \boldsymbol{W}^{1,2}, \tag{4.22}$$

(4.17) and (4.22) prove that vectors w'_k, $k = 1, \ldots, N$, belong to $\boldsymbol{D}(\boldsymbol{A}^*)$. ∎

Remark 4.5. Note that a key idea for the previous theorem is the choice of a type of not uniformly distributed splines.

Remark 4.6. In the case of just one pure delay, vectors generating subspaces Φ_n and Φ'_n become, respectively,

$$v_{jN+k} = \begin{bmatrix} 0 \\ \boldsymbol{\phi}_k \mathit{spline}_j \end{bmatrix}, \quad k = 1, \dots, N, \quad j = 1, 2, \dots, n, \tag{4.23}$$

$$v_k = \begin{bmatrix} \boldsymbol{\phi}_k \\ \boldsymbol{\phi}_k \mathit{spline}_0 \end{bmatrix}, \quad k = 1, \dots, N \tag{4.24}$$

for Φ_n, and

$$w_k^1 = \begin{bmatrix} 0 \\ \boldsymbol{\phi}_k \mathit{spline}_0 \end{bmatrix}, \quad k = 1, \dots, N, \tag{4.25}$$

$$w_{jN+k} = v_{jN+k} \qquad k = 1, \dots, N, \qquad j = 1, 2, \dots, n-1, \tag{4.26}$$

$$w'_k = \begin{bmatrix} \boldsymbol{\phi}_k \\ a_{1k} \mathit{spline}_n \end{bmatrix}, \quad k = 1, \dots, N \tag{4.27}$$

for Φ'_n. As can be seen, a great simplification is obtained with respect to the general case. Vectors v are just the ones in [3], and vectors w differ just for the fact that the nonzero term in $\mathbb{R}^N$ is taken from the first N vectors to the last ones, and the L_2 part of these last N vectors is multiplied for the columns of matrix $\boldsymbol{A}_1^{\mathrm{T}}$. This simplification with respect to the general case is due to the much simpler domain (2.21).

Consider now a test sequence of multi-indexes $\{s_n\}$, and consider the associated sequence of orthoprojection operators $\boldsymbol{\Pi}_{\Phi_n} : \boldsymbol{M}_2 \mapsto \Phi_n$ and $\boldsymbol{\Pi}_{\Phi'_n} : \boldsymbol{M}_2 \mapsto \Phi'_n$. For brevity, from now on the following notation is used:

$$\boldsymbol{\Pi}_n = \boldsymbol{\Pi}_{\Phi_n}, \qquad \boldsymbol{\Pi}'_n = \boldsymbol{\Pi}_{\Phi'_n}. \tag{4.28}$$

Recall that operators $\boldsymbol{\Pi}_n$ and $\boldsymbol{\Pi}'_n$, being orthogonal projectors, have the following properties:

$$\forall \boldsymbol{y} \in \boldsymbol{M}_2, \quad \|\boldsymbol{\Pi}_n \boldsymbol{y} - \boldsymbol{y}\| \le \|\boldsymbol{x} - \boldsymbol{y}\| \quad \forall \boldsymbol{x} \in \Phi_n, \tag{4.29}$$

$$\forall \boldsymbol{y} \in \boldsymbol{M}_2, \quad \|\boldsymbol{\Pi}'_n \boldsymbol{y} - \boldsymbol{y}\| \le \|\boldsymbol{x} - \boldsymbol{y}\| \quad \forall \boldsymbol{x} \in \Phi'_n. \tag{4.30}$$

The following results can be given on the convergence of the sequences of projectors $\boldsymbol{\Pi}_n$, $\boldsymbol{\Pi}'_n$, and of the sequence of semigroups generated by $\boldsymbol{\Pi}_n \boldsymbol{A} \boldsymbol{\Pi}_n$ and $\boldsymbol{\Pi}'_n \boldsymbol{A}^* \boldsymbol{\Pi}'_n$.

THEOREM 4.7. *The sequence of orthoprojection operators $\boldsymbol{\Pi}_n : \boldsymbol{M}_2 \mapsto \Phi_n$ converges strongly to the identity operator.*

Proof. Let $D = \{ \left[\begin{smallmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{smallmatrix}\right] \in \boldsymbol{M}_2 | \boldsymbol{y}_0 = \boldsymbol{y}_1(0), \boldsymbol{y}_1 \in C^2([-r, 0]; \mathbb{R}^N) \}$. Such set D is dense in $\boldsymbol{M}_2$ (see the proof of Lemma 2.2 and Remark 3.2 in [3]). Let $\boldsymbol{x} = \left[\begin{smallmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \end{smallmatrix}\right] \in D$ and let

$$\begin{aligned} \boldsymbol{x}^n = \begin{bmatrix} \boldsymbol{x}_0^n \\ \boldsymbol{x}_1^n \end{bmatrix} &= \sum_{k=1}^{N} (\boldsymbol{x}_1(-r)^{\mathrm{T}} \phi_k) v_{n_\delta N+k}^{\delta} + \sum_{k=1}^{N} \sum_{i=1}^{\delta-1} (\boldsymbol{x}_1(-r_i)^{\mathrm{T}} \phi_k) v_k^{r_i} \\ &+ \sum_{i=1}^{\delta} \sum_{k=1}^{N} \sum_{j=1}^{n_i-1} (\boldsymbol{x}_1(t_j^i)^{\mathrm{T}} \phi_k) v_{jN+k}^i + \sum_{k=1}^{N} (\boldsymbol{x}_1(0)^{\mathrm{T}} \phi_k) v_k^1. \end{aligned} \tag{4.31}$$

By Theorem 2.5 in [42] it is $\|\boldsymbol{x}_1^n - \boldsymbol{x}_1\| \to 0$, and the thesis follows by

$$\|\boldsymbol{x}^n - \boldsymbol{x}\|_{\boldsymbol{M}_2} = \|\boldsymbol{x}_1^n - \boldsymbol{x}_1\|_{\boldsymbol{L}_2} \tag{4.32}$$

and by property (4.29). ∎

THEOREM 4.8. *The sequence of semigroups* $\boldsymbol{T}_{\Phi_n}$ *generated by the operators* $\boldsymbol{\Pi}_n \boldsymbol{A} \boldsymbol{\Pi}_n$ *converges strongly to the semigroup governing the system* (2.5), (2.6).

Proof. Let D be the set in the proof of the previous theorem. There exists $\lambda > 0$ such that $(\boldsymbol{A} - \lambda \boldsymbol{I})D$ is dense in $\boldsymbol{M}_2$ (see Lemma 2.2 in [3]). There exists α such that $(\boldsymbol{A} - \alpha \boldsymbol{I})$ e $(\boldsymbol{\Pi}_n \boldsymbol{A} \boldsymbol{\Pi}_n - \alpha \boldsymbol{I})$ are dissipative (see Lemma 2.3 and proof of Theorem 3.1 in [3]). Let $\boldsymbol{x} = \left[\begin{smallmatrix}\boldsymbol{x}_0\\ \boldsymbol{x}_1\end{smallmatrix}\right] \in D$. Let $\boldsymbol{\Pi}_n \boldsymbol{x} = \left[\begin{smallmatrix}(\boldsymbol{\Pi}_n \boldsymbol{x})_0\\ (\boldsymbol{\Pi}_n \boldsymbol{x})_1\end{smallmatrix}\right]$. From

$$(\boldsymbol{\Pi}_n \boldsymbol{x})_1(-r_i) = (\boldsymbol{\Pi}_n \boldsymbol{x})_1(-r_{i-1}) - \int_{-r_i}^{-r_{i-1}} \frac{d(\boldsymbol{\Pi}_n \boldsymbol{x})_1(\vartheta)}{d\vartheta} d\vartheta \tag{4.33}$$

and as $\|\frac{d(\boldsymbol{x}_1 - (\boldsymbol{\Pi}_n \boldsymbol{x})_1)}{d\vartheta}\| \to 0$, (see Theorem 4.1 in [3], and Theorems 1.5, 2.5 in [42]), it follows that $\|\boldsymbol{A}\boldsymbol{\Pi}_n \boldsymbol{x} - \boldsymbol{A}\boldsymbol{x}\| \to 0$. Take into account that $(\boldsymbol{\Pi}_n \boldsymbol{x})_1(0) = (\boldsymbol{\Pi}_n \boldsymbol{x})_0$ and that $\|(\boldsymbol{\Pi}_n \boldsymbol{x})_0 - \boldsymbol{x}_0\| \to 0$.

Thus the Trotter–Kato theorem hypotheses are satisfied ([38], Lemma 3.1 in [3]). ∎

As can be seen, the proofs of the above two theorems follow the same lines of the proofs in [3], developed for the case of first order splines uniformly distributed in the interval $[-r, 0]$.

LEMMA 4.9. *The subspace*

$$U = \left\{ \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{bmatrix} \middle| \begin{array}{ll} \boldsymbol{y}_0 \in \mathbb{R}^N, & \boldsymbol{y}_1(0) = \boldsymbol{y}_0, \\ \boldsymbol{y}_1 \in W^{1,2} & \boldsymbol{y}_1(-r_j) = \boldsymbol{A}_j^{\mathrm{T}} \boldsymbol{y}_0, \ j = 1, \dots, \delta \end{array} \right\} \tag{4.34}$$

is dense in $\boldsymbol{M}_2$.

Proof. As usual, let us prove density in $\mathbb{R}^N \times W^{1,2}$. Let $\boldsymbol{y} = \left[\begin{smallmatrix}\boldsymbol{y}_0\\ \boldsymbol{y}_1\end{smallmatrix}\right] \in \mathbb{R}^N \times W^{1,2}$. Let us define the following sequence of functions $f_n : [-r, 0] \to \mathbb{R}^N$, $n > \sup_{i=1,2,\dots,\delta} \frac{r}{r_i - r_{i-1}}$,

$$\begin{aligned} f_n(\vartheta) = {} & \left(\boldsymbol{y}_0 - \frac{n}{r}\vartheta \left(\boldsymbol{y}_1\left(\frac{-r}{n}\right) - \boldsymbol{y}_0\right)\right) \chi_{[\frac{-r}{n}, 0]} \\ & + \left(\boldsymbol{A}_\delta^{\mathrm{T}} \boldsymbol{y}_0 + \frac{n}{r}(\vartheta + r)\left(\boldsymbol{y}_1\left(-r + \frac{r}{n}\right) - \boldsymbol{A}_\delta^{\mathrm{T}} \boldsymbol{y}_0\right)\right) \chi_{[-r, -r + \frac{r}{n}]} \\ & + \sum_{i=1}^{\delta} \left(\boldsymbol{A}_i^{\mathrm{T}} \boldsymbol{y}_0 + \frac{n}{r}(\vartheta + r_i)\left(\boldsymbol{y}_1\left(-r_i + \frac{r}{n}\right) - \boldsymbol{A}_i^{\mathrm{T}} \boldsymbol{y}_0\right)\right) \chi_{[-r_i, -r_i + \frac{r}{n}]} \\ & \quad + \left(\boldsymbol{A}_i^{\mathrm{T}} \boldsymbol{y}_0 - \frac{n}{r}(\vartheta + r_i)\left(\boldsymbol{y}_1\left(-r_i - \frac{r}{n}\right) - \boldsymbol{A}_i^{\mathrm{T}} \boldsymbol{y}_0\right)\right) \chi_{[-r_i - \frac{r}{n}, -r_i)}. \end{aligned} \tag{4.35}$$

Consider the sequence of elements in U,

$$\boldsymbol{y}_n = \begin{bmatrix} \boldsymbol{y}_0 \\ f_n + \sum_{i=1}^{\delta} \boldsymbol{y}_1 \chi_{(-r_i + \frac{r}{n}, -r_{i-1} - \frac{r}{n})} \end{bmatrix}. \tag{4.36}$$

As $\boldsymbol{y}_1$ is bounded, f_n is bounded too, uniformly on n. It follows that

$$\|\boldsymbol{y}_n - \boldsymbol{y}\|^2 \le \left(\sup_{\vartheta, n} \|\boldsymbol{y}_1(\vartheta) - f_n(\vartheta)\|^2\right) \frac{2\delta r}{n}. \tag{4.37}$$ ∎

Remark 4.10. The previous lemma proves that the intersection between the domain of $\boldsymbol{A}$ and the domain of $\boldsymbol{A}^*$ is dense in $\boldsymbol{M}_2$ if the weighted inner product is used. See that the subspace U is contained in both the domains. It is a standard result that such an intersection is in general not dense if the usual inner product is used [11, 14, 24, 27, 43].

THEOREM 4.11. *The sequence of orthoprojection operators* $\boldsymbol{\Pi}'_n : \boldsymbol{M}_2 \mapsto \Phi'_n$ *converges strongly to the identity operator.*

Proof. It is sufficient to prove strong convergence in a dense subspace of $\boldsymbol{M}_2$. Therefore, consider the subspace U in (4.34).

It is shown below that for any $\boldsymbol{y} = \left[\begin{smallmatrix}\boldsymbol{y}_0\\ \boldsymbol{y}_1\end{smallmatrix}\right] \in U$ there exists a sequence of approximations $\boldsymbol{y}^n \in \Phi'_n$ such that $\lim_{n\to\infty} \|\boldsymbol{y}^n - \boldsymbol{y}\|_{\boldsymbol{M}_2} = 0$. Consider the following definition of $\boldsymbol{y}^n \in \Phi'_n$:

$$
\begin{aligned}
\boldsymbol{y}^n &= \sum_{k=1}^{N} (\boldsymbol{y}_0^{\mathrm{T}}\phi_k) w'_k + \sum_{k=1}^{N}\sum_{i=1}^{\delta-1} (\boldsymbol{y}_1(-r_i)^{\mathrm{T}}\phi_k) w_k^{r_i},\\
&\quad + \sum_{i=1}^{\delta}\sum_{k=1}^{N}\sum_{j=1}^{n_i-1} (\boldsymbol{y}_1(t_j^i)^{\mathrm{T}}\phi_k) w^i_{jN+k} + \sum_{k=1}^{N} (\boldsymbol{y}_1(0)^{\mathrm{T}}\phi_k) w_k^1.
\end{aligned}
\tag{4.38}
$$

It is, by substituting expressions of vectors generating the subspace Φ'_n (4.6), (4.7), (4.8), (4.11),

$$
\begin{aligned}
\boldsymbol{y}^n = \begin{bmatrix}\boldsymbol{y}_0^n\\ \boldsymbol{y}_1^n\end{bmatrix} &= \begin{bmatrix}\boldsymbol{y}_0\\ \displaystyle\sum_{k=1}^{N} (\boldsymbol{y}_1(-r)^{\mathrm{T}}\phi_k)\phi_k\, spline^{\delta}_{n_\delta}\end{bmatrix}\\
&+ \begin{bmatrix}0\\ \displaystyle\sum_{k=1}^{N}\sum_{j=1}^{\delta-1}\left(\frac{(\boldsymbol{y}_0^{\mathrm{T}}\phi_k)a_{jk}}{\delta-j+1} spline''_{r_j} + (\boldsymbol{y}_1(-r_j)^{\mathrm{T}}\phi_k)\phi_k\, spline'_{r_j}\right)\end{bmatrix}\\
&+ \begin{bmatrix}0\\ \displaystyle\sum_{i=1}^{\delta}\sum_{k=1}^{N}\sum_{j=1}^{n_i-1} (\boldsymbol{y}_1(t_j^i)^{\mathrm{T}}\phi_k)\phi_k\, spline^i_j\end{bmatrix} + \begin{bmatrix}0\\ \displaystyle\sum_{k=1}^{N} (\boldsymbol{y}_1(0)^{\mathrm{T}}\phi_k)\phi_k\, spline^1_0\end{bmatrix}.
\end{aligned}
\tag{4.39}
$$

Moreover, it is readily recognized that

$$
\sum_{k=1}^{N} (\boldsymbol{y}_0^{\mathrm{T}}\phi_k) a_{AK} = \boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{y}_0 = \sum_{k=1}^{N} (\boldsymbol{y}_1(-r_j)^{\mathrm{T}}\phi_k)\phi_k, \quad j = 1,\ldots,\delta,
\tag{4.40}
$$

$$
\frac{1}{\delta-j+1} spline''_{r_j} + spline'_{r_j} = spline_{r_j}, \quad j = 1,\ldots,\delta-1,
\tag{4.41}
$$

so that

$$
\begin{aligned}
\|\boldsymbol{y}^n - \boldsymbol{y}\|_{\boldsymbol{M}_2} = \Bigg\| &\sum_{k=1}^{N} (\boldsymbol{y}_1(-r)^{\mathrm{T}}\phi_k)\phi_k\, spline^{\delta}_{n_\delta}\\
&+ \sum_{j=1}^{\delta-1} \boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{y}_0 \Big(\frac{1}{\delta-j+1} spline''_{r_j} + spline'_{r_j}\Big)
\end{aligned}
$$

$$+\sum_{i=1}^{\delta}\sum_{k=1}^{N}\sum_{j=1}^{n_i-1}(\boldsymbol{y}_1(t_j^i)^{\mathrm{T}}\phi_k)\phi_k \,spline_j^i$$

$$+\sum_{k=1}^{N}(\boldsymbol{y}_1(0)^{\mathrm{T}}\phi_k)\phi_k \,spline_0^1 - \boldsymbol{y}_1\Big\|_{L_2}$$

$$(4.42)\qquad = \Big\| \sum_{k=1}^{N}(\boldsymbol{y}_1(-r)^{\mathrm{T}}\phi_k)\phi_k \,spline_{n_\delta}^{\delta} + \sum_{j=1}^{\delta-1}\boldsymbol{y}_1(-r_j)\,spline_{r_j}$$

$$+\sum_{i=1}^{\delta}\sum_{k=1}^{N}\sum_{j=1}^{n_i-1}(\boldsymbol{y}_1(t_j^i)^{\mathrm{T}}\phi_k)\phi_k \,spline_j^i$$

$$+\sum_{k=1}^{N}(\boldsymbol{y}_1(0)^{\mathrm{T}}\phi_k)\phi_k \,spline_0^1 - \boldsymbol{y}_1\Big\|_{L_2}$$

$$= \Big\| \sum_{i=1}^{\delta}\sum_{k=1}^{N}\sum_{j=0}^{n_i}(\boldsymbol{y}_1(t_j^i)^{\mathrm{T}}\phi_k)\phi_k \,spline_j^i - \boldsymbol{y}_1\Big\|_{L_2}$$

$$= \sum_{i=1}^{\delta}\Big\| \sum_{k=1}^{N}\sum_{j=0}^{n_i}(\boldsymbol{y}_1(t_j^i)^{\mathrm{T}}\phi_k)\phi_k \,spline_j^i - \boldsymbol{y}_1\cdot\chi_{[-r_i,-r_{i-1}]}\Big\|_{L_2},$$

which gives the norm of the error between a function $\boldsymbol{y}_1 \in W^{1,2}$ and its approximation with first order splines in which the value at each spline knot (the instants t_j^i) is exactly the value of the function at time t_j^i. It is a standard result that the error tends to zero in L_2 norm for $n \to \infty$ (Theorem 2.4 in [42]) and therefore $\lim_{n\to\infty}\|\boldsymbol{y}^n - \boldsymbol{y}\|_{\boldsymbol{M}_2} = 0$. This implies, by property (4.30), the strong convergence to identity of operator $\boldsymbol{\Pi}'_n$. ∎

LEMMA 4.12. *There exists a real constant* α *such that the operator* $\boldsymbol{A}^* - \alpha\boldsymbol{I}$ *and operators* $\boldsymbol{\Pi}'_n\boldsymbol{A}^*\boldsymbol{\Pi}'_n - \alpha\boldsymbol{I}$ *are dissipative.*

Proof. In [3] it has been proved that there exists α such that operator $\boldsymbol{A} - \alpha\boldsymbol{I}$ is dissipative and therefore generates a semigroup which is a contraction one. This implies that the adjoint semigroup is a contraction one too and therefore its infinitesimal generator $\boldsymbol{A}^* - \alpha\boldsymbol{I}$ is dissipative [1]. Dissipativity of $\boldsymbol{A}^* - \alpha\boldsymbol{I}$ implies that for any n the operator $\boldsymbol{\Pi}'_n\boldsymbol{A}^*\boldsymbol{\Pi}'_n - \alpha\boldsymbol{I}$ is dissipative. This happens because for any $\boldsymbol{x} \in \boldsymbol{M}_2$

$$\big((\boldsymbol{\Pi}'_n\boldsymbol{A}^*\boldsymbol{\Pi}'_n - \alpha\boldsymbol{I})\boldsymbol{x},\boldsymbol{x}\big) = (\boldsymbol{A}^*\boldsymbol{\Pi}'_n\boldsymbol{x},\boldsymbol{\Pi}'_n\boldsymbol{x}) - \alpha(\boldsymbol{x},\boldsymbol{x})$$
$$\le (\boldsymbol{A}^*\boldsymbol{\Pi}'_n\boldsymbol{x},\boldsymbol{\Pi}'_n\boldsymbol{x}) - \alpha(\boldsymbol{\Pi}'_n\boldsymbol{x},\boldsymbol{\Pi}'_n\boldsymbol{x})$$
$$(4.43)\qquad = \big((\boldsymbol{A}^* - \alpha\boldsymbol{I})\boldsymbol{\Pi}'_n\boldsymbol{x},\boldsymbol{\Pi}'_n\boldsymbol{x}\big) \le 0. \qquad ∎$$

LEMMA 4.13. *Let* $\boldsymbol{D}$ *be the dense subspace of* $\boldsymbol{M}_2$ *defined as*

$$(4.44)\qquad \boldsymbol{D} = \left\{ \begin{bmatrix}\boldsymbol{y}_0\\ \boldsymbol{y}_1\end{bmatrix} \;\middle|\; \begin{matrix} \boldsymbol{y}_0 \in \mathbb{R}^N, \quad \boldsymbol{A}_\delta^{\mathrm{T}}\boldsymbol{y}_0 = \boldsymbol{y}_1(-r), \\ \left(\boldsymbol{y}_1 - \sum_{j=1}^{\delta-1}\boldsymbol{k}_j(\boldsymbol{y}_0,\boldsymbol{y}_1)\chi_{[-r,-r_j]}\right) \in C^2 \end{matrix} \right\}.$$

Then there exists $\lambda > 0$ *such that* $(\boldsymbol{A}^* - \lambda\boldsymbol{I})\boldsymbol{D}$ *is dense in* $\boldsymbol{M}_2$.

Proof. Let us at first assume the following additional property on the term $\boldsymbol{A}_{01}$ in the definition of operator $\boldsymbol{A}^*$:

$$(\mathrm{Hp}_0): \quad \frac{1}{g}\boldsymbol{A}_{01}^{\mathrm{T}} \quad \text{is a matrix of functions in } C^1([-r,0];\mathbb{R}),$$

where g is the weighting function in the inner product (2.4).

Hypothesis (Hp_0) will be removed at the end of the proof.

First, it will be shown that under assumption (Hp_0) there exists a sufficiently large λ and matrices $P_0^j(\lambda) \in \mathbb{R}^{N\times N}$ and $P_1^j(\lambda) \in \mathbb{R}^{N\times \delta N}$, $j = 1, 2, \ldots, \delta-1$, such that for any $\boldsymbol{z} = \left[\begin{smallmatrix}\boldsymbol{z}_0\\ \boldsymbol{z}_1\end{smallmatrix}\right] \in \mathbb{R}^N \times C^1([-r,0];\mathbb{R}^N)$ there exists $\boldsymbol{y} \in \boldsymbol{D}$ such that

$$(4.45) \qquad \begin{gathered} (\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y} = \begin{bmatrix} \boldsymbol{z}_0 \\ \tilde{\boldsymbol{z}}_1(\boldsymbol{z};\lambda)\end{bmatrix}, \\ \text{where} \quad \tilde{\boldsymbol{z}}_1(\boldsymbol{z};\lambda) = \boldsymbol{z}_1 + \sum_{j=1}^{\delta-1}\big(P_0^j(\lambda)\boldsymbol{z}_0 + P_1^j(\lambda)F_\lambda(\boldsymbol{z}_1)\big)\chi_{[-r,-r_j]}, \end{gathered}$$

in which the linear functional $F_\lambda(\boldsymbol{z}_1): C^1([-r,0];\mathbb{R}^N) \mapsto \mathbb{R}^{N\delta}$ is defined as follows:

$$(4.46) \qquad F_\lambda(\boldsymbol{z}_1) = \begin{pmatrix} \displaystyle\int_{-r}^{0} e^{\lambda\tau}\boldsymbol{z}_1(\tau)d\tau \\ \displaystyle\int_{-r}^{-r_1} e^{-\lambda(-r_1-\tau)}\boldsymbol{z}_1(\tau)d\tau \\ \vdots \\ \displaystyle\int_{-r}^{-r_{\delta-1}} e^{-\lambda(-r_{\delta-1}-\tau)}\boldsymbol{z}_1(\tau)d\tau \end{pmatrix}.$$

Next, it will be shown that there exists a sufficiently large λ such that for any given $\boldsymbol{x} = \left[\begin{smallmatrix}\boldsymbol{x}_0\\ \boldsymbol{x}_1\end{smallmatrix}\right] \in \boldsymbol{M}_2$ and for any $\varepsilon > 0$ there exists $\boldsymbol{z} = \left[\begin{smallmatrix}\boldsymbol{z}_0\\ \boldsymbol{z}_1\end{smallmatrix}\right] \in \mathbb{R}^N \times C^1([-r,0];\mathbb{R}^N)$ such that

$$(4.47) \qquad \left\| \begin{bmatrix}\boldsymbol{x}_0\\ \boldsymbol{x}_1\end{bmatrix} - \begin{bmatrix}\boldsymbol{z}_0\\ \tilde{\boldsymbol{z}}_1(\boldsymbol{z};\lambda)\end{bmatrix}\right\|_{\boldsymbol{M}_2} \le \varepsilon$$

and therefore, from (4.45),

$$(4.48) \qquad \forall \boldsymbol{x} \in \boldsymbol{M}_2,\ \forall\varepsilon > 0, \quad \exists \boldsymbol{y} \in \boldsymbol{D}: \ \big\|\boldsymbol{x} - (\boldsymbol{A}^* - \lambda\boldsymbol{I})\boldsymbol{y}\big\|_{\boldsymbol{M}_2} \le \varepsilon,$$

that is, the density of $(\boldsymbol{A}^* - \lambda\boldsymbol{I})\boldsymbol{D}$ in $\boldsymbol{M}_2$.

In order to prove (4.45) as a first step it is shown how to find a function $\bar{Y}_1(\boldsymbol{y}_0, \boldsymbol{z}_1)$ such that for any $\boldsymbol{y}_0 \in \mathbb{R}^N$ and $\boldsymbol{z}_1 \in C^1([-r,0];\mathbb{R}^N)$ it is $\left[\begin{smallmatrix}\boldsymbol{y}_0\\ \bar{Y}_1(\boldsymbol{y}_0,\boldsymbol{z}_1)\end{smallmatrix}\right] \in \boldsymbol{D}$. Next, it is shown how to define a function $Y_0(\boldsymbol{z}_0, \boldsymbol{z}_1): \mathbb{R}^N \times C^1([-r,0];\mathbb{R}^N) \to \mathbb{R}^N$ such that the composed function $Y_1(\boldsymbol{z}_0,\boldsymbol{z}_1) = \bar{Y}_1(Y_0(\boldsymbol{z}_0,\boldsymbol{z}_1),\boldsymbol{z}_1)$ has the property

$$(4.49) \qquad (\boldsymbol{A}^* - \lambda\boldsymbol{I})\begin{bmatrix} Y_0(\boldsymbol{z}_0,\boldsymbol{z}_1)\\ Y_1(\boldsymbol{z}_0,\boldsymbol{z}_1)\end{bmatrix} = \begin{bmatrix}\boldsymbol{z}_0\\ \tilde{\boldsymbol{z}}_1(\boldsymbol{z}_1;\lambda)\end{bmatrix}.$$

For any given pair $\boldsymbol{y}_0 \in \mathbb{R}^N$ and $\boldsymbol{z}_1 \in C^1([-r,0];\mathbb{R}^N)$ let us consider the differential equation in $C^2([-r,0];\mathbb{R}^N)$,

$$(4.50) \qquad \frac{d}{d\vartheta}\boldsymbol{f}(\vartheta) + \lambda\boldsymbol{f}(\vartheta) = -\Big(\boldsymbol{z}_1(\vartheta) - \frac{1}{g(\vartheta)}\boldsymbol{A}_{01}^{\mathrm{T}}(\vartheta)\boldsymbol{y}_0\Big),$$

whose solution is

$$\boldsymbol{f}(\vartheta) = e^{-\lambda(\vartheta+r)}\boldsymbol{f}(-r) - \int_{-r}^{\vartheta} e^{-\lambda(\vartheta-\tau)}\Big(\boldsymbol{z}_1(\tau) - \frac{1}{g(\tau)}\boldsymbol{A}_{01}^{\mathrm{T}}(\tau)\boldsymbol{y}_0\Big)d\tau. \tag{4.51}$$

By Lemma A.1 in the appendix, there exists a unique left-continuous function $\boldsymbol{y}_1$ that satisfies condition (A.1), with $\boldsymbol{f}$ given by (4.51). Such $\boldsymbol{y}_1$ is given by expression (A.6), and its values at the delay instants are such that

$$(\boldsymbol{I}_{(\delta-1)N} - H_{\delta,2})\begin{bmatrix} \boldsymbol{y}_1(-r_1) \\ \vdots \\ \boldsymbol{y}_1(-r_{\delta-1}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}(-r_1) \\ \vdots \\ \boldsymbol{f}(-r_{\delta-1}) \end{bmatrix} - H_{\delta,2}\begin{bmatrix} \boldsymbol{A}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{A}_{\delta-1}^{\mathrm{T}} \end{bmatrix}\boldsymbol{y}_0, \tag{4.52}$$

where $H_{\delta,2}$ is defined in (A.4), Lemma A.1, in the appendix.

In order to guarantee that $\boldsymbol{y} = \left[\begin{smallmatrix}\boldsymbol{y}_0\\ \boldsymbol{y}_1\end{smallmatrix}\right] \in \boldsymbol{D}$, $\boldsymbol{y}_1$ must satisfy the additional condition

$$\boldsymbol{y}_1(-r) = \boldsymbol{A}_\delta^{\mathrm{T}}\boldsymbol{y}_0. \tag{4.53}$$

By substituting (4.53) in (A.1) one has

$$\boldsymbol{f}(-r) = \boldsymbol{A}_\delta^{\mathrm{T}}\boldsymbol{y}_0 - \sum_{j=1}^{\delta-1}\frac{\boldsymbol{y}_1(-r_j) - \boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{y}_0}{\delta - j + 1}, \tag{4.54}$$

which can be rewritten as

$$\boldsymbol{f}(-r) = h_{\delta,1}\begin{bmatrix} \boldsymbol{y}_1(-r_1) \\ \vdots \\ \boldsymbol{y}_1(-r_{\delta-1}) \end{bmatrix} - h_{\delta,2}\begin{bmatrix} \boldsymbol{A}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{A}_{\delta-1}^{\mathrm{T}} \end{bmatrix}\boldsymbol{y}_0, \tag{4.55}$$

where

$$\begin{aligned} h_{\delta,1} &= \begin{bmatrix} \frac{1}{\delta}\boldsymbol{I}_N & \cdots & \frac{1}{2}\boldsymbol{I}_N & \boldsymbol{I}_N \end{bmatrix}, \\ h_{\delta,2} &= \begin{bmatrix} \frac{1}{\delta}\boldsymbol{I}_N & \cdots & \frac{1}{2}\boldsymbol{I}_N \end{bmatrix}. \end{aligned} \tag{4.56}$$

By (4.51) the values of function $\boldsymbol{f}$ at the delay instants are as follows:

$$\begin{bmatrix} \boldsymbol{f}(-r_1) \\ \vdots \\ \boldsymbol{f}(-r_{\delta-1}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_N e^{-\lambda(r-r_1)} \\ \vdots \\ \boldsymbol{I}_N e^{-\lambda(r-r_{\delta-1})} \end{bmatrix} \boldsymbol{f}(-r) \quad - \begin{bmatrix} 0_{(\delta-1)N\times N} & \boldsymbol{I}_{(\delta-1)N} \end{bmatrix} F_\lambda\left(\boldsymbol{z}_1 - \frac{1}{g}A_{01}^{\mathrm{T}}\boldsymbol{y}_0\right). \tag{4.57}$$

By substituting (4.52) and (4.55) into (4.57) and rearranging we have

$$H_p(\lambda)\begin{bmatrix} \boldsymbol{y}_1(-r_1) \\ \vdots \\ \boldsymbol{y}_1(-r_{\delta-1}) \end{bmatrix} = H_q(\lambda)\begin{bmatrix} \boldsymbol{A}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{A}_{\delta-1}^{\mathrm{T}} \end{bmatrix} \boldsymbol{y}_0 - \begin{bmatrix} 0_{(\delta-1)N\times N} & \boldsymbol{I}_{(\delta-1)N} \end{bmatrix}\left(F_\lambda(\boldsymbol{z}_1) - F_\lambda\left(\frac{1}{g}A_{01}^{\mathrm{T}}\right)\boldsymbol{y}_0\right), \tag{4.58}$$

in which matrices H_p and H_q are defined as

$$H_p(\lambda) = \boldsymbol{I}_{(\delta-1)N} - H_{\delta,2} + \begin{bmatrix} \boldsymbol{I}_N e^{-\lambda(r-r_1)} \\ \vdots \\ \boldsymbol{I}_N e^{-\lambda(r-r_{\delta-1})} \end{bmatrix} h_{\delta,2}, \tag{4.59}$$

$$H_q(\lambda) = \begin{bmatrix} \boldsymbol{I}_N e^{-\lambda(r-r_1)} \\ \vdots \\ \boldsymbol{I}_N e^{-\lambda(r-r_{\delta-1})} \end{bmatrix} h_{\delta,1} - \begin{bmatrix} H_{\delta,2} \\ 0_{N\times\delta N} \end{bmatrix}. \tag{4.60}$$

Because H_p is nonsingular (Lemma A.3 in the appendix), by (4.58) it results that

$$\begin{aligned} \begin{bmatrix} \boldsymbol{y}_1(-r_1) \\ \vdots \\ \boldsymbol{y}_1(-r_{\delta-1}) \end{bmatrix} &= \left(H_p^{-1}(\lambda) H_q(\lambda) \begin{bmatrix} \boldsymbol{A}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{A}_{\delta-1}^{\mathrm{T}} \end{bmatrix} + \left[0_{(\delta-1)N\times N} \; H_p^{-1}(\lambda) \right] F_\lambda \left(\frac{1}{g} \boldsymbol{A}_{01}^{\mathrm{T}} \right) \right) \boldsymbol{y}_0 \\ &\quad - \left[0_{(\delta-1)N\times N} \; H_p^{-1}(\lambda) \right] F_\lambda(\boldsymbol{z}_1). \end{aligned} \tag{4.61}$$

From (4.61) and (4.53), recalling that $r_\delta = r$, matrices $N_j(\lambda)$ and $M_j(\lambda)$, $j = 1, \dots, \delta$ are defined such that

$$\boldsymbol{y}_1(-r_j) = N_j(\lambda)\boldsymbol{y}_0 + M_j(\lambda) F_\lambda(\boldsymbol{z}_1). \tag{4.62}$$

The left-continuous function $\boldsymbol{y}_1 = \bar{Y}_1(\boldsymbol{y}_0, \boldsymbol{z}_1)$ we were looking for is given by

$$\boldsymbol{y}_1(\vartheta) = \begin{cases} \boldsymbol{y}_1(-r_i), & \vartheta = -r_i, \\ \boldsymbol{f}(\vartheta) + \displaystyle\sum_{j=1}^{\delta-1} \frac{\boldsymbol{y}_1(-r_j) - \boldsymbol{A}_j^{\mathrm{T}} \boldsymbol{y}_0}{\delta - j + 1} \chi_{[-r,-r_j]}(\vartheta), & \vartheta \neq -r_i, \end{cases} \quad i = 1, \dots, \delta - 1, \tag{4.63}$$

in which $\boldsymbol{f}(\vartheta)$ is given by (4.51). This is such that $\left[\begin{smallmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{smallmatrix}\right] \in \boldsymbol{D}$. Let $(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y} = \left[\begin{smallmatrix} [(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}]_0 \\ [(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}]_1 \end{smallmatrix}\right]$. It is

$$\begin{aligned} [(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}]_1 &= \frac{1}{g} \boldsymbol{A}_{01}^{\mathrm{T}} \boldsymbol{y}_0 - \frac{d}{d\vartheta} \left(\boldsymbol{y}_1 - \sum_{j=1}^{\delta-1} \frac{\boldsymbol{y}_1(-r_j) - \boldsymbol{A}_j^{\mathrm{T}} \boldsymbol{y}_0}{\delta - j + 1} \chi_{[-r,-r_j]} \right) - \lambda \boldsymbol{y}_1 \\ &= \frac{1}{g} \boldsymbol{A}_{01}^{\mathrm{T}} \boldsymbol{y}_0 - \frac{d}{d\vartheta} \boldsymbol{f} - \lambda \boldsymbol{f} - \lambda \sum_{j=1}^{\delta-1} \frac{\boldsymbol{y}_1(-r_j) - \boldsymbol{A}_j^{\mathrm{T}} \boldsymbol{y}_0}{\delta - j + 1} \chi_{[-r,-r_j]}. \end{aligned} \tag{4.64}$$

Finally, recalling the definition (4.50) of function $\boldsymbol{f}$, it is

$$[(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}]_1 = \boldsymbol{z}_1 - \lambda \sum_{j=1}^{\delta-1} \frac{\left(N_j(\lambda) - \boldsymbol{A}_j^{\mathrm{T}}\right)\boldsymbol{y}_0 + M_j(\lambda) F_\lambda(\boldsymbol{z}_1)}{\delta - j + 1} \chi_{[-r,-r_j]}. \tag{4.65}$$

Until now we have showed that, for any $\boldsymbol{y}_0 \in \mathbb{R}^N$ and for any $\boldsymbol{z}_1 \in C^1([-r,0];\mathbb{R}^N)$ it is possible to find $\boldsymbol{y}_1$ such that $\boldsymbol{y} = \left[\begin{smallmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{smallmatrix}\right] \in \boldsymbol{D}$ and satisfies (4.65).

Now, using the computed $\boldsymbol{y}_1 = \bar{Y}_1(\boldsymbol{y}_0, \boldsymbol{z}_1)$, we are ready to prove that there exists a positive λ such that for any $\boldsymbol{z}_0 \in \mathbb{R}^N$ and $\boldsymbol{z}_1 \in C^1([-r,0];\mathbb{R}^N)$, a $\boldsymbol{y}_0$ can be found such that $\boldsymbol{y} = \left[\begin{smallmatrix}\boldsymbol{y}_0\\ \boldsymbol{y}_1\end{smallmatrix}\right] \in \boldsymbol{D}$ and satisfies (4.45). The application of operator $(\boldsymbol{A}^* - \lambda \boldsymbol{I})$ gives, for the part in $\mathbb{R}^N$,

$$[(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}]_0 = \delta\, \boldsymbol{y}_1(0) + \boldsymbol{A}_0^{\mathrm{T}} \boldsymbol{y}_0 - \lambda \boldsymbol{y}_0. \tag{4.66}$$

Note that from (4.63) $\boldsymbol{y}_1(0) = \boldsymbol{f}(0)$ and evaluation of $\boldsymbol{f}(0)$ according to (4.51), in which expression (4.54) of $\boldsymbol{f}(-r)$ is substituted, gives

$$[(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}]_0 = Q_0(\lambda)\boldsymbol{y}_0 + Q_1(\lambda) F_\lambda(\boldsymbol{z}_1), \tag{4.67}$$

in which

$$\begin{aligned} Q_0(\lambda) &= \delta e^{-\lambda r} \boldsymbol{A}_\delta^{\mathrm{T}} - \delta e^{-\lambda r} \sum_{j=1}^{\delta-1} \frac{N_j(\lambda) - \boldsymbol{A}_j^{\mathrm{T}}}{(\delta - j + 1)} + \delta \int_{-r}^{0} e^{\lambda \tau} \frac{1}{g} \boldsymbol{A}_{01}^{\mathrm{T}} d\tau + \boldsymbol{A}_0^{\mathrm{T}} - \lambda \boldsymbol{I}_N, \\ Q_1(\lambda) &= -\delta \left[\, \boldsymbol{I}_{N\times N} 0_{N \times N(\delta-1)} \,\right] - \sum_{j=1}^{\delta-1} \frac{\delta e^{-\lambda r}}{\delta - j + 1} M_j(\lambda). \end{aligned} \tag{4.68}$$

It is clear that there exists a sufficiently large λ such that $Q_0(\lambda)$ is nonsingular, due to the presence of the term $-\lambda \boldsymbol{I}_N$ (the other terms are all bounded functions of λ). Therefore, given $\boldsymbol{z}_0 \in \mathbb{R}^N$ and $\boldsymbol{z}_1 \in C^1([-r,0];\mathbb{R}^N)$, the function $Y_0(\boldsymbol{z}_0, \boldsymbol{z}_1) = \boldsymbol{y}_0 \in \mathbb{R}^N$ such that $[(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}]_0 = \boldsymbol{z}_0$, thanks to (4.67), is given by

$$\boldsymbol{y}_0 = Y_0(\boldsymbol{z}_0, \boldsymbol{z}_1) = Q_0^{-1}(\lambda)(\boldsymbol{z}_0 - Q_1(\lambda) F_\lambda(\boldsymbol{z}_1)). \tag{4.69}$$

Substitution of (4.69) in the expression (4.65) for $[(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}]_1$ gives

$$\begin{aligned} [(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}]_1 = \boldsymbol{z}_1 &- \lambda \sum_{j=1}^{\delta-1} \frac{\left(N_j(\lambda) - \boldsymbol{A}_j^{\mathrm{T}}\right) Q_0^{-1}(\lambda) \boldsymbol{z}_0}{(\delta - j + 1)} \chi_{[-r,-r_j]} \\ &- \lambda \sum_{j=1}^{\delta-1} \frac{\left(N_j(\lambda) - \boldsymbol{A}_j^{\mathrm{T}}\right) Q_0^{-1}(\lambda) Q_1(\lambda) F_\lambda(\boldsymbol{z}_1) + M_j(\lambda) F_\lambda(\boldsymbol{z}_1)}{(\delta - j + 1)} \chi_{[-r,-r_j]}. \end{aligned} \tag{4.70}$$

This expression allows one to define the matrices $P_0^j(\lambda)$ and $P_1^j(\lambda)$ used in (4.45) as

$$\begin{aligned} P_0^j &= -\lambda \frac{\left(N_j(\lambda) - \boldsymbol{A}_j^{\mathrm{T}}\right) Q_0^{-1}(\lambda)}{(\delta - j + 1)}, \\ P_1^j &= -\lambda \frac{\left(N_j(\lambda) - \boldsymbol{A}_j^{\mathrm{T}}\right) Q_0^{-1}(\lambda) Q_1(\lambda) + M_j(\lambda)}{(\delta - j + 1)}. \end{aligned} \tag{4.71}$$

Composition of functions $\bar{Y}_1(\boldsymbol{y}_0, \boldsymbol{z}_1)$ and $Y_0(\boldsymbol{z}_0, \boldsymbol{z}_1)$ gives the announced function $Y_1(\boldsymbol{z}_0, \boldsymbol{z}_1)$. This concludes the proof that, for λ sufficiently large, for any $\left[\begin{smallmatrix}\boldsymbol{z}_0\\ \boldsymbol{z}_1\end{smallmatrix}\right] \in \mathbb{R}^N \times C^1([-r,0];\mathbb{R}^N)$ there exists $\left[\begin{smallmatrix}\boldsymbol{y}_0\\ \boldsymbol{y}_1\end{smallmatrix}\right] \in \boldsymbol{D}$ such that (4.45) holds.

Consider now the continuous linear function Φ_λ defined as follows:

$$\begin{aligned} &\Phi^{(\lambda)} : L_2([-r,0];\mathbb{R}^N) \mapsto L_2([-r,0];\mathbb{R}^N), \\ &\Phi^{(\lambda)}(\boldsymbol{g}) = \boldsymbol{g} + \sum_{j=1}^{\delta-1} P_1^j(\lambda) F_\lambda(\boldsymbol{g}) \chi_{[-r,-r_j]}. \end{aligned} \tag{4.72}$$

Let us define the following subspace of $L_2([-r,0];\mathbb{R}^N)$:

$$\mathcal{R} = \Phi^{(\lambda)}(C^1([-r,0];\mathbb{R}^N)). \tag{4.73}$$

The proof of the lemma is obtained if the set $\mathcal{R}$ is proved to be dense in $L_2([-r,0];\mathbb{R}^N)$. This is true because it can be readily shown that density of $\mathcal{R}$ is sufficient to conclude that $\forall \boldsymbol{x} \in \boldsymbol{M}_2$, for any $\varepsilon > 0$, there exists a $\boldsymbol{y} \in \boldsymbol{D}$ such that $\|\boldsymbol{x} - (\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}\|_{\boldsymbol{M}_2} \leq \varepsilon$ (i.e., density of $(\boldsymbol{A}^* - \lambda\boldsymbol{I})\boldsymbol{D}$).

Given a $\boldsymbol{x} = \left[\begin{smallmatrix}\boldsymbol{x}_0\\ \boldsymbol{x}_1\end{smallmatrix}\right] \in \boldsymbol{M}_2$, take $\boldsymbol{y}_A \in \boldsymbol{D}$ as follows: $\boldsymbol{y}_{A,0} = Y_0(\boldsymbol{x}_0, 0) = Q_0^{-1}(\lambda)\boldsymbol{x}_0$ and $\boldsymbol{y}_{A,1} = Y_1(\boldsymbol{x}_0, 0)$. It is, by construction,

$$(\boldsymbol{A}^* - \lambda\boldsymbol{I})\boldsymbol{y}_A = \begin{bmatrix} \boldsymbol{x}_0 \\ \sum_{j=1}^{\delta-1} P_0^j(\lambda)\boldsymbol{x}_0\chi_{[-r,-r_j]} \end{bmatrix}. \tag{4.74}$$

From the density of $\mathcal{R}$, there exists $\boldsymbol{z}_{B,1} \in C^1([-r,0];\mathbb{R}^N)$ such that the function

$$\tilde{\boldsymbol{z}}_{B,1} = \boldsymbol{z}_{B,1} + \sum_{j=1}^{\delta-1} P_1^j(\lambda)F_\lambda(\boldsymbol{z}_{B,1})\chi_{[-r,-r_j]} \tag{4.75}$$

satisfies

$$\left\| \sum_{j=1}^{\delta-1} P_0^j(\lambda)\boldsymbol{x}_0\chi_{[-r,-r_j]} - \tilde{\boldsymbol{z}}_{B,1} \right\|_{L_2} \leq \frac{\varepsilon}{2}, \tag{4.76}$$

and from result (4.45) there exists $\boldsymbol{y}_B \in \boldsymbol{D}$ such that $(\boldsymbol{A}^* - \lambda\boldsymbol{I})\boldsymbol{y}_B = \left[\begin{smallmatrix} 0 \\ \tilde{\boldsymbol{z}}_{B,1}\end{smallmatrix}\right]$.

Exploiting again the density of $\mathcal{R}$ there exists $\boldsymbol{z}_{C,1} \in C^1([-r,0];\mathbb{R}^N)$ such that the function

$$\tilde{\boldsymbol{z}}_{C,1} = \boldsymbol{z}_{C,1} + \sum_{j=1}^{\delta-1} P_1^j(\lambda)F_\lambda(\boldsymbol{z}_{C,1})\chi_{[-r,-r_j]} \tag{4.77}$$

satisfies

$$\left\| \boldsymbol{x}_1 - \tilde{\boldsymbol{z}}_{C,1} \right\|_{L_2} \leq \frac{\varepsilon}{2}. \tag{4.78}$$

Again, from result (4.45) there exists $\boldsymbol{y}_C \in \boldsymbol{D}$ such that $(\boldsymbol{A}^* - \lambda\boldsymbol{I})\boldsymbol{y}_C = \left[\begin{smallmatrix} 0 \\ \tilde{\boldsymbol{z}}_{C,1}\end{smallmatrix}\right]$. It is now an easy matter to show that vector $\boldsymbol{y} = \boldsymbol{y}_A - \boldsymbol{y}_B + \boldsymbol{y}_C$ is such that

$$\begin{aligned} \|\boldsymbol{x} - (A^* - \lambda\boldsymbol{I})\boldsymbol{y}\|_{M_2} &= \|\boldsymbol{x} - (A^* - \lambda\boldsymbol{I})(\boldsymbol{y}_A - \boldsymbol{y}_B + \boldsymbol{y}_C)\|_{\boldsymbol{M}_2} \\ &\leq \left\| \begin{bmatrix} 0 \\ \boldsymbol{x}_1 - [(A^* - \lambda\boldsymbol{I})\boldsymbol{y}_A]_1 + \tilde{\boldsymbol{z}}_{B,1} - \tilde{\boldsymbol{z}}_{C,1} \end{bmatrix} \right\|_{\boldsymbol{M}_2} \\ &\leq \left\|\boldsymbol{x}_1 - \tilde{\boldsymbol{z}}_{C,1}\right\|_{L_2} + \left\| \sum_{j=1}^{\delta-1} P_0^j(\lambda)\boldsymbol{x}_0\chi_{[-r,-r_j]} - \tilde{\boldsymbol{z}}_{B,1} \right\|_{L_2} \leq \varepsilon. \end{aligned} \tag{4.79}$$

It remains to prove that $\mathcal{R}$ is dense for sufficiently large λ. We will show that if for any $\boldsymbol{f} \in L_2([-r,0];\mathbb{R}^N)$ there exists a vector $\alpha \in \mathbb{R}^{N\delta}$ such that

$$F_\lambda(\boldsymbol{f}) - F_\lambda\left(\sum_{j=1}^{\delta-1} P_1^j \alpha \chi_{[-r,-r_j]} \right) - \alpha = 0, \tag{4.80}$$

then for any $\boldsymbol{f} \in L_2([-r,0];\mathbb{R}^N)$ a sequence $\{\boldsymbol{f}_k\}$, $\boldsymbol{f}_k \in \mathcal{R}\ \forall k \geq 0$ can be found such that $\|\boldsymbol{f} - \boldsymbol{f}_k\|_{L_2} \to 0$. Existence of α in (4.80) for any $\boldsymbol{f}$ is ensured by the nonsingularity of matrix

$$\Gamma(\lambda) = \boldsymbol{I}_{N\delta\times N\delta} + \sum_{j=1}^{\delta-1} F_\lambda\big(P_1^j(\lambda)\chi_{[-r,-r_j]}\big) \tag{4.81}$$

for sufficiently large λ, and this is a sufficient condition for density of $\mathcal{R}$.

To this purpose consider a $\boldsymbol{f} \in L_2$, let α be the solution of (4.80), and define the function

$$\bar{\boldsymbol{f}} = \boldsymbol{f} - \sum_{j=1}^{\delta-1} P_1^j \alpha \chi_{[-r,-r_j]}. \tag{4.82}$$

It is such that $\Phi^{(\lambda)}(\bar{\boldsymbol{f}}) = \boldsymbol{f}$. Let $\{\boldsymbol{g}_k\}$ be a sequence of functions in $C^1([-r,0];\mathbb{R}^N)$ such that $\|\bar{\boldsymbol{f}} - \boldsymbol{g}_k\|_{L_2} \to 0$. From the continuity of function $\Phi^{(\lambda)}$ it is $\|\Phi^{(\lambda)}(\bar{\boldsymbol{f}}) - \Phi^{(\lambda)}(\boldsymbol{g}_k)\|_{L_2} \to 0$. Defining functions $\boldsymbol{f}_k = \Phi^{(\lambda)}(\boldsymbol{g}_k) \in \mathcal{R}$, the sequence $\{\boldsymbol{f}_k\}$ converges to $\Phi^{(\lambda)}(\bar{\boldsymbol{f}})$, that is, $\boldsymbol{f}$ and density of $\mathcal{R}$, under nonsingularity of $\Gamma(\lambda)$, is proved.

It remains to prove the nonsingularity of the $\delta N \times \delta N$ matrix $\Gamma(\lambda)$ defined in (4.81) for a sufficiently large λ. Such a proof is reported in [39] and is worked out by showing that $\det\big(\Gamma(\lambda)\big)$ is a continuous function of λ and that there exists the limit matrix $\bar{\Gamma} = \lim_{\lambda\to+\infty}\Gamma(\lambda)$. Such a matrix can be easily proved to be nonsingular, because it is block triangular (each block is $N \times N$), in which the diagonal consists of the following nonsingular δ blocks: block 1 is $\boldsymbol{I}_N$, block j, for $j = 2,\ldots,\delta$, is $\boldsymbol{I} + \frac{1}{\delta-j+1}\boldsymbol{I}_N$. It follows that $\lim_{\lambda\to+\infty}\det\big(\Gamma(\lambda)\big) = \det\big(\bar{\Gamma}\big) \neq 0$, and therefore there exists λ_0 such that for every $\lambda > \lambda_0$ matrix $\Gamma(\lambda)$ is nonsingular.

So, chosen λ such that $\Gamma(\lambda)$ and $Q_0(\lambda)$ are both nonsingular, the proof of this lemma is completed in the case of hypothesis Hp_0.

To remove such a hypothesis it is sufficient to consider a sequence $\boldsymbol{A}_{01}^k$ in the space $L_2([-r,0];\mathbb{R}^{N\times N})$, which converges to $\boldsymbol{A}_{01}$ and satisfies hypothesis Hp_0. Let $\boldsymbol{A}_k^*$ be the corresponding sequence of operators. Let $D_k \in L(\boldsymbol{M}_2)$ be defined as

$$D_k \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{g}(\boldsymbol{A}_{01}^k - \boldsymbol{A}_{01})^{\mathrm{T}}\boldsymbol{y}_0 \end{bmatrix}.$$

Thus $\boldsymbol{A}_k^* = \boldsymbol{A}^* + D_k$ and $\|D_k\|_{L(\boldsymbol{M}_2)} \leq \|\boldsymbol{A}_{01}^k - \boldsymbol{A}_{01}\|_{L_2([-r,0];\mathbb{R}^{N\times N})}$. From Proposition 2.3 in [5, page 28], and Theorem 1.1 in [38, page 76], it follows that any λ with $Re(\lambda) > \omega_0 + \sup_k \|D_k\|$ belongs to the resolvent set of $\boldsymbol{A}^*$ and $\boldsymbol{A}_k^*$, for every k, where ω_0 is such that $\|T(t)\| \leq e^{\omega_0 t}$, $\boldsymbol{T}(t)$ being the semigroup generated by $\boldsymbol{A}$ (see Lemma 2.3 in [3]). Let us choose a λ in the resolvent sets of $\boldsymbol{A}^*$ and $\boldsymbol{A}_k^*$, such that $(\boldsymbol{A}_k^* - \lambda\boldsymbol{I})\boldsymbol{D}$ is dense in $\boldsymbol{M}_2$ for every k. It is sufficient that corresponding matrices $\Gamma^k(\lambda)$ in (4.81) and $Q_0^k(\lambda)$ in (4.68) are nonsingular. Thus, given $\boldsymbol{x} \in M_2$, given $\epsilon > 0$, a sequence $\boldsymbol{y}_k$ can be found such that

$$\|(\boldsymbol{A}_k^* - \lambda\boldsymbol{I})\boldsymbol{y}_k - \boldsymbol{x}\| < \frac{\epsilon}{2}. \tag{4.83}$$

From

$$\|(\boldsymbol{A}^* - \lambda\boldsymbol{I})\boldsymbol{y}_k - \boldsymbol{x}\| \leq \|(\boldsymbol{A}_k^* - \lambda\boldsymbol{I})\boldsymbol{y}_k - \boldsymbol{x}\| + \|D_k\|\|\boldsymbol{y}_k\| \tag{4.84}$$

it follows that there exist k_0 such that $\|(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{y}_{k_0} - \boldsymbol{x}\| < \epsilon$, provided $\boldsymbol{y}_k$ is uniformly bounded. It is sufficient that $\|D_{k_0}\| < \frac{\epsilon}{2\sup_k \|\boldsymbol{y}_k\|}$. It remains to prove uniform boundedness of $\boldsymbol{y}_k$. Let $v_k = (\boldsymbol{A}_k^* - \lambda \boldsymbol{I})\boldsymbol{y}_k - \boldsymbol{x}$. From (4.83) it is $\|v_k\| < \frac{\epsilon}{2}$ for every k. From

$$(\boldsymbol{A}_k^* - \lambda \boldsymbol{I})^{-1} = (\boldsymbol{A}^* + D_k - \lambda \boldsymbol{I})^{-1} = [\boldsymbol{I} - (\lambda \boldsymbol{I} - \boldsymbol{A}^*)^{-1} D_k]^{-1} (\boldsymbol{A}^* - \lambda \boldsymbol{I})^{-1} \tag{4.85}$$

it follows that

$$\|(\boldsymbol{A}_k^* - \lambda \boldsymbol{I})^{-1}\| \leq \|[\boldsymbol{I} - (\lambda \boldsymbol{I} - \boldsymbol{A}^*)^{-1} D_k]^{-1}\| \|(\boldsymbol{A}^* - \lambda \boldsymbol{I})^{-1}\|. \tag{4.86}$$

If k is sufficiently large such that $\|(\lambda \boldsymbol{I} - \boldsymbol{A}^*)^{-1} D_k\| \leq d < 1$, the following inequality holds:

$$\|[\boldsymbol{I} - (\lambda \boldsymbol{I} - \boldsymbol{A}^*)^{-1} D_k]^{-1}\| \leq \sum_{l=0}^{\infty} \|(\lambda \boldsymbol{I} - \boldsymbol{A}^*)^{-1} D_k\|^l \leq \frac{1}{1-d}, \tag{4.87}$$

which proves the uniform boundedness of $\|(\boldsymbol{A}_k^* - \lambda \boldsymbol{I})^{-1}\|$. The uniform boundedness of $\boldsymbol{y}_k$ follows by

$$\boldsymbol{y}_k = (\boldsymbol{A}_k^* - \lambda \boldsymbol{I})^{-1} (v_k - \boldsymbol{x}). \tag{4.88}$$

Such a device to prove the density of $(\boldsymbol{A}^* - \lambda \boldsymbol{I})\boldsymbol{D}$ in $\boldsymbol{M}_2$ when hypothesis Hp_0 is not satisfied has been introduced in [24, Theorem 7.2] for the one delay case. ◻

LEMMA 4.14. *The operator* $\boldsymbol{\Pi}_n' \boldsymbol{A}^* \boldsymbol{\Pi}_n'$ *converges strongly to the operator* $\boldsymbol{A}^*$ *in the subspace* $\boldsymbol{D}$ *defined in Lemma* 4.13.

Proof. Since it is

$$\|\boldsymbol{\Pi}_n' \boldsymbol{A}^* \boldsymbol{\Pi}_n' \boldsymbol{x} - \boldsymbol{A}^* \boldsymbol{x}\|_{M_2} \leq \|\boldsymbol{\Pi}_n' (\boldsymbol{A}^* \boldsymbol{\Pi}_n' - \boldsymbol{A}^*) \boldsymbol{x}\|_{M_2} + \|\boldsymbol{\Pi}_n' \boldsymbol{A}^* \boldsymbol{x} - \boldsymbol{A}^* \boldsymbol{x}\|_{M_2}, \tag{4.89}$$

and $\lim_{n\to\infty} \|\boldsymbol{\Pi}_n' \boldsymbol{y} - \boldsymbol{y}\|_{M_2} = 0$, $\forall \boldsymbol{y} \in \boldsymbol{M}_2$ (strong convergence), the lemma is proved if for any $\boldsymbol{x} \in \boldsymbol{D}$

$$\|\boldsymbol{A}^* \boldsymbol{\Pi}_n' \boldsymbol{x} - \boldsymbol{A}^* \boldsymbol{x}\| \to 0. \tag{4.90}$$

It is

$$\begin{aligned}
\|\boldsymbol{A}^* \boldsymbol{\Pi}_n' \boldsymbol{x} - \boldsymbol{A}^* \boldsymbol{x}\|^2 &= \|\delta \boldsymbol{x}_1(0) + \boldsymbol{A}_0^{\mathrm{T}} \boldsymbol{x}_0 - \delta(\boldsymbol{\Pi}_n' \boldsymbol{x})_1(0) - \boldsymbol{A}_0^{\mathrm{T}} (\boldsymbol{\Pi}_n' \boldsymbol{x})_0\|^2 \\
&\quad + \Bigg\| \frac{1}{g} \boldsymbol{A}_{01} \boldsymbol{x}_0 - \frac{1}{g} \boldsymbol{A}_{01} (\boldsymbol{\Pi}_n' \boldsymbol{x})_0 - \frac{d}{d\vartheta} \left(\boldsymbol{x}_1 - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{x}_0, \boldsymbol{x}_1) \chi_{[-r,-r_j]} \right) \\
&\quad + \frac{d}{d\vartheta} \left((\boldsymbol{\Pi}_n' \boldsymbol{x})_1 - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j((\boldsymbol{\Pi}_n' \boldsymbol{x})_0, (\boldsymbol{\Pi}_n' \boldsymbol{x})_1) \chi_{[-r,-r_j]} \right) \Bigg\|_{L_2}^2 \\
&\leq \delta^2 \|\boldsymbol{x}_1(0) - (\boldsymbol{\Pi}_n' \boldsymbol{x})_1(0)\|^2 + \left(\|\boldsymbol{A}_0^{\mathrm{T}}\| + 2 \left\| \frac{1}{g} \boldsymbol{A}_{01}^{\mathrm{T}} \right\|_{L_2} \right) \cdot \left\| \boldsymbol{x}_0 - (\boldsymbol{\Pi}_n' \boldsymbol{x})_0 \right\|^2 + 2 S_n^2(\boldsymbol{x}),
\end{aligned} \tag{4.91}$$

where

$$\begin{aligned}
S_n(\boldsymbol{x}) = \Bigg\| &\frac{d}{d\vartheta} \left(\boldsymbol{x}_1 - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{x}_0, \boldsymbol{x}_1) \chi_{[-r,-r_j]} \right) \\
&- \frac{d}{d\vartheta} \left((\boldsymbol{\Pi}_n' \boldsymbol{x})_1 - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j((\boldsymbol{\Pi}_n' \boldsymbol{x})_0, (\boldsymbol{\Pi}_n' \boldsymbol{x})_1) \chi_{[-r,-r_j]} \right) \Bigg\|_{L_2}.
\end{aligned} \tag{4.92}$$

Strong convergence of $\mathbf{\Pi}'_n$ ensures that for $n \to \infty$ the second term in the right-hand side of (4.91) goes to zero.

To prove that the term $S_n(\boldsymbol{x})$ goes to zero too, let

$$\begin{aligned}\bar{\boldsymbol{x}} = \begin{bmatrix}\bar{\boldsymbol{x}}_0\\ \bar{\boldsymbol{x}}_1\end{bmatrix} &= \sum_{k=1}^{N}\boldsymbol{x}_0(k)w'_k + \sum_{i=1}^{\delta-1}\sum_{k=1}^{N}(\boldsymbol{x}_1(-r_i)^{\mathrm{T}}\phi_k)w_k^{r_i}\\ &\quad+\sum_{i=1}^{\delta}\sum_{j=1}^{n_i}\sum_{k=1}^{N}(\boldsymbol{x}_1(t_j^i)^{\mathrm{T}}\phi_k)w^i_{jN+k} + \sum_{k=1}^{N}(\boldsymbol{x}_1(0)^{\mathrm{T}}\phi_k)w_k^1.\end{aligned} \tag{4.93}$$

It is such that $\bar{\boldsymbol{x}}_0 = \boldsymbol{x}_0$, so that $\|\boldsymbol{x}-\bar{\boldsymbol{x}}\|_{\boldsymbol{M}_2} = \|\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1\|_{L_2}$ and therefore

$$\|\boldsymbol{x} - \mathbf{\Pi}'_n\boldsymbol{x}\|_{\boldsymbol{M}_2} \le \|\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1\|_{L_2}. \tag{4.94}$$

Considering that the function $\sum_{j=1}^{\delta-1}\boldsymbol{k}_j(\cdot,\cdot)\chi_{[-r,-r_j]}$ is piecewise constant it is

$$S_n(\boldsymbol{x}) = \left\|\frac{d}{d\vartheta}\boldsymbol{x}_1 - \frac{d}{d\vartheta}(\mathbf{\Pi}'_n\boldsymbol{x})_1\right\|_{L_2} \tag{4.95}$$

and

$$S_n(\boldsymbol{x}) \le \left\|\frac{d}{d\vartheta}\boldsymbol{x}_1 - \frac{d}{d\vartheta}\bar{\boldsymbol{x}}_1\right\|_{L_2} + \left\|\frac{d}{d\vartheta}\bar{\boldsymbol{x}}_1 - \frac{d}{d\vartheta}(\mathbf{\Pi}'_n\boldsymbol{x})_1\right\|_{L_2}. \tag{4.96}$$

As for the first term at the right-hand side of inequality (4.96), since it is

$$\left\|\frac{d}{d\vartheta}\boldsymbol{x}_1 - \frac{d}{d\vartheta}\bar{\boldsymbol{x}}_1\right\|_{L_2} = \left(\sum_{i=1}^{\delta}\left\|\frac{d}{d\vartheta}\boldsymbol{x}_1\chi_{[-r_i,-r_{i-1}]} - \frac{d}{d\vartheta}\bar{\boldsymbol{x}}_1\chi_{[-r_i,-r_{i-1}]}\right\|^2_{L_2}\right)^{\frac{1}{2}}, \tag{4.97}$$

by standard results of spline analysis (see Theorem 2.5 in [42]), each term in the summation goes to zero for $n \to \infty$.

As for the second term, from definition of vectors w that generate V'_n, it is, by applying the Schmidt inequality (see Theorem 1.5 in [42]),

$$\begin{aligned}\left\|\frac{d}{d\vartheta}\bar{\boldsymbol{x}}_1 - \frac{d}{d\vartheta}(\mathbf{\Pi}'_n\boldsymbol{x})_1\right\|_{L_2} &= \left(\sum_{i=1}^{\delta}\left\|\frac{d}{d\vartheta}\bar{\boldsymbol{x}}_1\chi_{[-r_i,-r_{i-1}]} - \frac{d}{d\vartheta}(\mathbf{\Pi}'_n\boldsymbol{x})_1\chi_{[-r_i,-r_{i-1}]}\right\|^2_{L_2}\right)^{\frac{1}{2}}\\ &\le \left(\sum_{i=1}^{\delta}\left(\sqrt{12}\frac{n_i}{r_i - r_{i-1}}\left\|\left(\bar{\boldsymbol{x}}_1 - (\mathbf{\Pi}'_n\boldsymbol{x})_1\right)\chi_{[-r_i,-r_{i-1}]}\right\|_{L_2}\right)^2\right)^{\frac{1}{2}}.\end{aligned}$$

For each term in the summation it is

$$\begin{aligned}\left\|\left(\bar{\boldsymbol{x}}_1 - (\mathbf{\Pi}'_n\boldsymbol{x})_1\right)\chi_{[-r_i,-r_{i-1}]}\right\|_{L_2} &\le \left\|\bar{\boldsymbol{x}}_1 - (\mathbf{\Pi}'_n\boldsymbol{x})_1\right\|_{L_2}\\ &\le \left\|\bar{\boldsymbol{x}}_1 - \boldsymbol{x}_1\right\|_{L_2} + \left\|\boldsymbol{x}_1 - (\mathbf{\Pi}'_n\boldsymbol{x})_1\right\|_{L_2} \le 2\|\bar{\boldsymbol{x}}_1 - \boldsymbol{x}_1\|\\ &\le \left(\sum_{i=1}^{\delta}\left(\left\|(\bar{\boldsymbol{x}}_1 - \boldsymbol{x}_1)\chi_{[-r_i,-r_{i-1}]}\right\|_{L_2}\right)^2\right)^{\frac{1}{2}}.\end{aligned} \tag{4.98}$$

Again, by standard results on spline approximation (Theorem 2.5 in [42]), each term in the summation goes to zero for $n \to \infty$. This proves that $S_n(\boldsymbol{x})$ goes to zero for $n \to \infty$.

It remains to prove that the term $\delta\|\boldsymbol{x}_1(0) - (\boldsymbol{\Pi}'_n\boldsymbol{x})_1(0)\|^2$ in the right-hand side of (4.91) goes to zero for $n \to \infty$.

First, note that being $\boldsymbol{x} \in \mathcal{D}(\boldsymbol{A}^*)$ it is such that for $i = 1, \ldots, \delta - 1$

$$\boldsymbol{x}_1(-r_i^+) - \sum_{j=1}^{i-1} \boldsymbol{k}_j(\boldsymbol{x}_0, \boldsymbol{x}_1) = \boldsymbol{x}_1(-r_i) - \sum_{j=1}^{i} \boldsymbol{k}_j(\boldsymbol{x}_0, \boldsymbol{x}_1), \tag{4.99}$$

where $\boldsymbol{x}_1(-r_i^+)$ denotes the limit of $\boldsymbol{x}_1(\vartheta)$ for ϑ approaching $-r_i$ from the right (note that in general $\boldsymbol{x}_1(-r_i^+) \neq \boldsymbol{x}_1(-r_i)$). Simple computations, taking into account definition (2.22) of $\boldsymbol{k}_j$, give

$$\boldsymbol{x}_1(-r_i^+) = \frac{\delta - i + 2}{\delta - i + 1}\boldsymbol{x}_1(-r_i) - \frac{1}{\delta - i + 1}\boldsymbol{A}_i^{\mathrm{T}}\boldsymbol{x}_0. \tag{4.100}$$

Since also $\boldsymbol{\Pi}'_n\boldsymbol{x} \in \mathcal{D}(\boldsymbol{A}^*)$, it is such that

$$(\boldsymbol{\Pi}'_n\boldsymbol{x})_1(-r_i^+) = \frac{\delta - i + 2}{\delta - i + 1}(\boldsymbol{\Pi}'_n\boldsymbol{x})_1(-r_i) - \frac{1}{\delta - i + 1}\boldsymbol{A}_i^{\mathrm{T}}(\boldsymbol{\Pi}'_n\boldsymbol{x})_0. \tag{4.101}$$

At point $-r$ it is

$$\boldsymbol{x}_1(-r) = \boldsymbol{A}_\delta^{\mathrm{T}}\boldsymbol{x}_0, \qquad (\boldsymbol{\Pi}'_n\boldsymbol{x})_1(-r) = \boldsymbol{A}_\delta^{\mathrm{T}}(\boldsymbol{\Pi}'_n\boldsymbol{x})_0. \tag{4.102}$$

Since it has been proved that $\lim_{n\to\infty}\|\boldsymbol{x}_0 - (\boldsymbol{\Pi}'_n\boldsymbol{x})_0\| = 0$, from (4.102) it follows

$$\lim_{n\to\infty}\|\boldsymbol{x}_1(-r) - (\boldsymbol{\Pi}'_n\boldsymbol{x})_1(-r)\| = 0. \tag{4.103}$$

Starting from (4.103), the proof that $\|\boldsymbol{x}_1(0) - (\boldsymbol{\Pi}'_n\boldsymbol{x})_1(0)\|$ goes to zero is obtained if we prove the following recursive implication:

$$\lim_{n\to\infty}\|\boldsymbol{x}_1(-r_i) - (\boldsymbol{\Pi}'_n\boldsymbol{x})_1(-r_i)\| = 0, \tag{4.104}$$

$$\Downarrow$$

$$\lim_{n\to\infty}\|\boldsymbol{x}_1(-r_{i-1}) - (\boldsymbol{\Pi}'_n\boldsymbol{x})_1(-r_{i-1})\| = 0. \tag{4.105}$$

First note that if (4.104) is true, then comparing (4.100) and (4.101), recalling that $\|\boldsymbol{x}_0 - (\boldsymbol{\Pi}'_n\boldsymbol{x})_0\| \to 0$, it follows

$$\lim_{n\to\infty}\|\boldsymbol{x}_1(-r_i^+) - (\boldsymbol{\Pi}'_n\boldsymbol{x})_1(-r_i^+)\| = 0. \tag{4.106}$$

From (4.106), since it has been proved that in any interval $[-r_i, -r_{i-1}]$

$$\lim_{n\to\infty}\left\|\left(\frac{d}{d\vartheta}\boldsymbol{x}_1 - \frac{d}{d\vartheta}(\boldsymbol{\Pi}'_n\boldsymbol{x})_1\right)\chi_{[-r_i, -r_{i-1}]}\right\|_{L_2} = 0, \tag{4.107}$$

implication (4.105) is easily obtained. This completes the proof of the Lemma. ∎

THEOREM 4.15. *The sequence of semigroups $\boldsymbol{T}^*_{\Phi'_n}$ generated by the operators $\boldsymbol{\Pi}'_n\boldsymbol{A}^*\boldsymbol{\Pi}'_n$ converges strongly to $\boldsymbol{T}^*$, the adjoint of the semigroup generated by $\boldsymbol{A}$.*

Proof. The results in Lemmas 4.12, 4.13, and 4.14 imply that the hypotheses of the Trotter–Kato theorem, as stated in [38] and reported also in Lemma 3.1 in [3], are satisfied, and this proves the convergence of $\boldsymbol{T}^*_{\Phi'_n}$ to $\boldsymbol{T}^*$. ∎

Now the main result of the paper can be given, that is, the theorem on the convergence of the proposed finite dimensional approximation scheme of the LQG controller for hereditary systems.

THEOREM 4.16. *Let* Φ_n *and* Φ'_n *be the sequences of finite dimension subspaces of* $\boldsymbol{M}_2$ *in Definitions* 4.2, 4.3. *Let* $\boldsymbol{u}_n(t)$ *be the input function obtained by*

$$\boldsymbol{u}_n(t) = -\boldsymbol{B}^*\boldsymbol{R}_n(t_f - t)\boldsymbol{\Pi}'_n\hat{\boldsymbol{x}}_n(t), \tag{4.108}$$

where

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}}_n(t) &= \boldsymbol{\Pi}_n\boldsymbol{A}\boldsymbol{\Pi}_n\hat{\boldsymbol{x}}_n(t) + \boldsymbol{\Pi}_n\boldsymbol{B}\boldsymbol{u}_n(t) + \boldsymbol{P}_n(t)\boldsymbol{\Pi}_n\boldsymbol{C}^*\big(\boldsymbol{y}(t) - \boldsymbol{C}\boldsymbol{\Pi}_n\hat{\boldsymbol{x}}_n(t)\big), \\ \hat{\boldsymbol{x}}_n(0) &= \boldsymbol{\Pi}_n\hat{\boldsymbol{x}}(0) \end{aligned} \tag{4.109}$$

in which $\boldsymbol{P}_n$ *and* $\boldsymbol{R}_n$ *are the finite dimensional solutions of the Riccati equations* (3.15) *and* (3.16) *in which the projectors* $\boldsymbol{\Pi}_n$ *and* $\boldsymbol{\Pi}'_n$ *are considered. Let* $\boldsymbol{u}(t)$ *be the optimal input,* $\hat{\boldsymbol{x}}(t)$ *the optimal estimated state,* $\boldsymbol{x}_n(t)$ *and* $\boldsymbol{x}(t)$ *the actual state evolving when* $\boldsymbol{u}_n(t)$ *and* $\boldsymbol{u}(t)$ *are applied to system* (2.1), (2.2), *respectively.*

Then

$$\lim_{n\to\infty} E\|\boldsymbol{x}_n - \boldsymbol{x}\|^2_{L_2([0,t_f];\boldsymbol{M}_2)} = 0, \tag{4.110}$$

$$\lim_{n\to\infty} E\|\hat{\boldsymbol{x}}_n - \hat{\boldsymbol{x}}\|^2_{L_2([0,t_f];\boldsymbol{M}_2)} = 0, \tag{4.111}$$

$$\lim_{n\to\infty} E\|\boldsymbol{u}_n - \boldsymbol{u}\|^2_{L_2([0,t_f];\mathbb{R}^p)} = 0, \tag{4.112}$$

$$\lim_{n\to\infty} |J_f(\boldsymbol{u}_n) - J_f(\boldsymbol{u})| = 0. \tag{4.113}$$

Proof. The proof comes from Theorem 3.7, whose assumptions (from Hp_1 to Hp_4) are satisfied thanks to Theorems 4.7, 4.8, 4.11, and 4.15. ∎

5. Implementation of the method. In this section the numerical implementation of the approximation scheme described in the previous section, and which satisfies all properties listed in the introduction, is reported.

Consider two Hilbert spaces $\mathcal{U}$ and $\mathcal{V}$ and two finite dimensional subspaces $U_n \subset \mathcal{U}$ and $V_m \subset \mathcal{V}$ of dimension n and m, respectively. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ be a basis of U_n and $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ a basis of V_m. Consider the nonsingular matrices $\boldsymbol{T}_n \in \mathbb{R}^{n\times n}$ and $Z_m \in \mathbb{R}^{m\times m}$, whose components are defined as

$$\begin{aligned} T_n(i,j) &= (\boldsymbol{u}_i, \boldsymbol{u}_j)_{\mathcal{U}}, \quad i,j = 1,\dots,n, \\ Z_m(h,k) &= (\boldsymbol{v}_i, \boldsymbol{v}_j)_{\mathcal{V}}, \quad i,j = 1,\dots,m. \end{aligned} \tag{5.1}$$

Recall that the orthoprojection operator from $\mathcal{U}$ to U_n performs the following operation on a vector $\boldsymbol{x} \in \mathcal{U}$:

$$\boldsymbol{\Pi}(\boldsymbol{x}; U_n) = \sum_{i=1}^{n} \alpha_i\boldsymbol{u}_i \quad \text{with} \quad \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = T_n^{-1} \begin{bmatrix} (\boldsymbol{x}, \boldsymbol{u}_1)_{\mathcal{U}} \\ \vdots \\ (\boldsymbol{x}, \boldsymbol{u}_n)_{\mathcal{U}} \end{bmatrix}, \tag{5.2}$$

and the orthoprojection operator from $\mathcal{V}$ to V_m performs the following operation on a vector $\boldsymbol{y} \in H_2$:

$$\boldsymbol{\Pi}(\boldsymbol{y}; V_m) = \sum_{i=1}^{m} \beta_i\boldsymbol{v}_i \quad \text{with} \quad \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} = Z_m^{-1} \begin{bmatrix} (\boldsymbol{y}, \boldsymbol{v}_1)_{\mathcal{V}} \\ \vdots \\ (\boldsymbol{y}, \boldsymbol{v}_m)_{\mathcal{V}} \end{bmatrix}. \tag{5.3}$$

Let us denote as ξ_n the isomorphism that associates to a vector $\boldsymbol{x} \in U_n$ its coordinate representation

$$\xi_n : \; U_n \mapsto \mathbb{R}^n; \quad \xi_n(\boldsymbol{x}) = T_n^{-1} \begin{bmatrix} (\boldsymbol{x}, \boldsymbol{u}_1)_{\mathcal{U}} \\ \vdots \\ (\boldsymbol{x}, \boldsymbol{u}_n)_{\mathcal{U}} \end{bmatrix} \tag{5.4}$$

and as ξ_m the isomorphism

$$\xi_m : \; V_m \mapsto \mathbb{R}^m; \quad \xi_m(\boldsymbol{y}) = Z_m^{-1} \begin{bmatrix} (\boldsymbol{y}, \boldsymbol{v}_1)_{\mathcal{V}} \\ \vdots \\ (\boldsymbol{y}, \boldsymbol{v}_m)_{\mathcal{V}} \end{bmatrix}. \tag{5.5}$$

Consider now the algebra $\mathcal{S}$ of linear operators from U_n to V_m. It is

$$S \in \mathcal{S}, \quad \xi_m(S(\boldsymbol{u}_i)) = Z_m^{-1} \begin{bmatrix} (S(\boldsymbol{u}_i), \boldsymbol{v}_1)_{\mathcal{V}} \\ \vdots \\ (S(\boldsymbol{u}_i), \boldsymbol{v}_m)_{\mathcal{V}} \end{bmatrix}. \tag{5.6}$$

The following isomorphism η_n^m is induced between $\mathcal{S}$ and the algebra of matrices $m \times n$:

$$S \in \mathcal{S}, \quad \eta_n^m(S) = Z_m^{-1} \bar{S}, \quad \bar{S} \in \mathbb{R}^{m \times n}, \quad \bar{S}_{i,j} = (S(\boldsymbol{u}_j), \boldsymbol{v}_i)_{\mathcal{V}}, \tag{5.7}$$

that is, such that

$$\xi_m(S(\boldsymbol{x})) = \eta_n^m(S)\, \xi_n(\boldsymbol{x}), \quad \boldsymbol{x} \in U_n. \tag{5.8}$$

Isomorphisms between points of finite dimensional spaces and their coordinate representations and between linear operators on finite dimensional spaces and their matrix representations allow us to write the approximated Riccati equations for control (3.16) and for filtering (3.15) as

$$\begin{aligned} \dot{\widetilde{\boldsymbol{P}}}_n(t) &= \widetilde{\boldsymbol{A}}_n \boldsymbol{W}_n^{-1} \widetilde{\boldsymbol{P}}_n(t) + \widetilde{\boldsymbol{P}}_n(t) \boldsymbol{W}_n^{-1} \widetilde{\boldsymbol{A}}_n^{\mathrm{T}} + \widetilde{\boldsymbol{\Lambda}}_n b - \widetilde{\boldsymbol{P}}_n(t) \boldsymbol{W}_n^{-1} \widetilde{\boldsymbol{\Sigma}}_n \boldsymbol{W}_n^{-1} \widetilde{\boldsymbol{P}}_n(t), \\ \dot{\widetilde{\boldsymbol{R}}}_n(t) &= \widetilde{\mathcal{A}}_n \mathcal{W}_n^{-1} \widetilde{\boldsymbol{R}}_n(t) + \widetilde{\boldsymbol{R}}_n(t) \mathcal{W}_n^{-1} \widetilde{\mathcal{A}}_n^{\mathrm{T}} + \widetilde{\boldsymbol{L}}_n - \widetilde{\boldsymbol{R}}_n(t) \mathcal{W}_n^{-1} \widetilde{\boldsymbol{S}}_n \mathcal{W}^{-1} \widetilde{\boldsymbol{R}}_n(t) \end{aligned} \tag{5.9}$$

and the approximated filter equation (3.46) and control equation (3.47) in the form

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}}_{n,C}(t) &= \boldsymbol{W}_n^{-1}([\widetilde{\boldsymbol{A}}_n - \widetilde{\boldsymbol{P}}_n(t)\widetilde{\boldsymbol{\Sigma}}_n]\hat{\boldsymbol{x}}_{n,C}(t) + \widetilde{\boldsymbol{P}}_n(t)\boldsymbol{W}_n^{-1}\bar{\Gamma}_n \boldsymbol{y}(t) - \widetilde{\boldsymbol{T}}_n(t_f - t)\hat{\boldsymbol{x}}_{n,C}(t)), \\ \boldsymbol{u}_n(t) &= -\begin{bmatrix} 0 & \boldsymbol{B}_0^T \end{bmatrix} \mathcal{W}_n^{-1} \widetilde{\boldsymbol{R}}_n(t_f - t) \mathcal{W}_n^{-1} \widetilde{\boldsymbol{T}}_{n,2} \hat{\boldsymbol{x}}_{n,C}(t), \\ \hat{\boldsymbol{z}}_n(t) &= [\, \boldsymbol{I}_{N \times N} \quad 0_{N \times nN} \,]\, \hat{\boldsymbol{x}}_{n,C}(t). \end{aligned} \tag{5.10}$$

In (5.9) and (5.10) $\hat{\boldsymbol{x}}_{n,C}$ is the coordinate expression of vector $\hat{\boldsymbol{x}}_n$ in the basis of Φ_n, $\hat{\boldsymbol{z}}_n(t)$ is the approximation of the optimal estimate of $\boldsymbol{z}(t)$, $\widetilde{\boldsymbol{P}}_n(t)$ and $\widetilde{\boldsymbol{R}}_n(t)$ are square matrices whose components are defined as $\{\widetilde{\boldsymbol{P}}_n(t)\}_{i,j} = (\boldsymbol{P}_n(t)\boldsymbol{v}_j, \boldsymbol{v}_i)$ and $\{\widetilde{\boldsymbol{R}}_n(t)\}_{i,j} = (\boldsymbol{R}_n(t)\boldsymbol{w}_j, \boldsymbol{w}_i)$. Matrices $\widetilde{\boldsymbol{\Lambda}}_n$, $\widetilde{\boldsymbol{\Sigma}}_n$, $\bar{\boldsymbol{\Gamma}}_n$, $\boldsymbol{W}_n$, $\widetilde{\mathcal{A}}_n$, $\widetilde{\boldsymbol{T}}_{n,1}$, $\mathcal{W}_n$, $\widetilde{\boldsymbol{T}}_{n,2}(t)$,

$\widetilde{\mathcal{A}}_n$, $\widetilde{\boldsymbol{L}}_n$, $\widetilde{\boldsymbol{S}}_n$ are numerical matrices computed by simple scalar products of elements in finite dimensional subspaces as follows:

$$
\begin{aligned}
&\widetilde{\boldsymbol{\Lambda}}_n(i,j) = (\boldsymbol{F}\boldsymbol{F}^*\boldsymbol{v}_j, \boldsymbol{v}_i)\,,\\
&\widetilde{\boldsymbol{\Sigma}}_n(i,j) = (\boldsymbol{C}^*\boldsymbol{C}\boldsymbol{v}_j, \boldsymbol{v}_i)\,,\\
&\bar{\boldsymbol{\Gamma}}_n(i,j) = (\boldsymbol{C}^*\phi_j, \boldsymbol{v}_i)\,,\\
&\boldsymbol{W}_n(i,j) = (\boldsymbol{v}_i, \boldsymbol{v}_j),\\
&\widetilde{\boldsymbol{A}}_n(i,j) = (\boldsymbol{A}\boldsymbol{v}_j, \boldsymbol{v}_i),\\
&\widetilde{\boldsymbol{T}}_{n,1}(i,j) = (\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{w}_j, \boldsymbol{v}_i)\,,\\
&\mathbb{W}_n(i,j) = (\boldsymbol{w}_i, \boldsymbol{w}_j),\\
&\widetilde{\boldsymbol{T}}_{n,2}(i,j) = (\boldsymbol{v}_j, \boldsymbol{w}_i),\\
&\widetilde{\mathcal{A}}_n(i,j) = (\boldsymbol{A}^*\boldsymbol{w}_j, \boldsymbol{w}_i)\,,\\
&\widetilde{\boldsymbol{L}}_n(i,j) = (\boldsymbol{Q}\boldsymbol{w}_j, \boldsymbol{w}_i),\\
&\widetilde{\boldsymbol{S}}_n(i,j) = (\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{w}_j, \boldsymbol{w}_i).
\end{aligned}
\tag{5.11}
$$

Finally, it is $\widetilde{\boldsymbol{T}}_n(t) = \widetilde{\boldsymbol{T}}_{n,1}\mathbb{W}_n^{-1}\widetilde{\boldsymbol{R}}_n(t)\mathbb{W}_n^{-1}\widetilde{\boldsymbol{T}}_{n,2}$.

Thus, denoting by

$$
\begin{aligned}
S_c(n,t) &= \boldsymbol{W}_n^{-1}(\widetilde{\boldsymbol{A}}_n - \widetilde{\boldsymbol{P}}_n(t)\boldsymbol{W}_n^{-1}\widetilde{\boldsymbol{\Sigma}}_n - \widetilde{\boldsymbol{T}}_{n,1}\mathbb{W}_n^{-1}\widetilde{\boldsymbol{R}}_n(t_f - t)\mathbb{W}_n^{-1}\widetilde{\boldsymbol{T}}_{n,2}),\\
P_c(n,t) &= \boldsymbol{W}_n^{-1}\widetilde{\boldsymbol{P}}_n(t)\boldsymbol{W}_n^{-1}\bar{\boldsymbol{\Gamma}}_n,\\
Q_c(n,t) &= -\left[\,0 \quad \boldsymbol{B}_0^{\mathrm{T}}\,\right]\mathbb{W}_n^{-1}\widetilde{\boldsymbol{R}}_n(t_f - t)\mathbb{W}_n^{-1}\widetilde{\boldsymbol{T}}_{n,2}
\end{aligned}
\tag{5.12}
$$

with $\widetilde{\boldsymbol{P}}_n(t)$ and $\widetilde{\boldsymbol{R}}_n(t)$ solutions of the matrix differential equations (Riccati) in (5.9), the approximate LQG controller can be written as follows:

$$
\begin{aligned}
\dot{\hat{\boldsymbol{x}}}_{n,C}(t) &= S_c(n,t)\hat{\boldsymbol{x}}_{n,C}(t) + P_c(n,t)\boldsymbol{y}(t),\\
\boldsymbol{u}_n(t) &= Q_c(n,t)\hat{\boldsymbol{x}}_{n,C}(t).
\end{aligned}
\tag{5.13}
$$

The vector $\hat{\boldsymbol{x}}_{n,C}(t) \in \mathbb{R}^{(n_1+1+\sum_{i=2}^{\delta} n_i)N}$.

Remark 5.1. It is important to stress the fact that matrices in (5.11) have a fixed structure and, in the case of hereditary systems without distributed delay, such matrices depend only on the multiindex s_n and on the matrices $\boldsymbol{A}_j$ $(j = 0, 1, \dots, \delta)$, $\boldsymbol{B}_0$, $\boldsymbol{C}_0$, $\boldsymbol{F}_0$, $\boldsymbol{G}$ that describe the system and on the weight matrix Q_0 that defines the cost functional. This property follows from the fact that splines are not uniformly distributed over the interval $[-r, 0]$: each interval $[-r_i, -r_{i-1}]$ has an independent spline distribution.

The numerical computation of matrices (5.11) is a straightforward function of the multi-index s_n and of the system matrices. As an example, the expressions of matrices in (5.11) are reported for systems with two pure delay terms (multi-index $s_n = (n_1, n_2)$ with $n = \inf(n_1, n_2)$) and no distributed delay.

$$
\boldsymbol{W}_n = \begin{bmatrix} \boldsymbol{W}_{n,a} & \boldsymbol{W}_{n,b} \\ 0_{n_2N\times n_1N} & \boldsymbol{W}_{n,c} \end{bmatrix},
\tag{5.14}
$$

$$\boldsymbol{W}_{n,a} = \left(\frac{r_1}{n_1}\right)\begin{bmatrix} \frac{n_1}{r_1}+2/3 & 1/3 & 0 & \dots & 0 \\ 1/3 & 4/3 & 1/3 & \ddots & \vdots \\ 0 & 1/3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 4/3 & 1/3 \\ 0 & \cdot & 0 & 1/3 & 2/3+\frac{n_1}{r_1}\frac{r-r_1}{3n_2} \end{bmatrix}_{(n_1+1)\times(n_1+1)} \otimes \boldsymbol{I}_{N\times N},$$

$$\boldsymbol{W}_{n,b} = \begin{bmatrix} 0 & \dots & \dots & 0 \\ \frac{r-r_1}{6n_2}\boldsymbol{I}_{N\times N} & 0 & \dots & 0 \end{bmatrix}_{(n_1+1)N\times n_2N},$$

$$\boldsymbol{W}_{n,c} = \left(\frac{r-r_1}{n_2}\right)\begin{bmatrix} 1/6 & 2/3 & 1/6 & 0 & \dots \\ 0 & 1/6 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2/3 & 1/6 \\ 0 & \cdot & 0 & 1/6 & 1/3 \end{bmatrix}_{n_2\times(n_2+1)} \otimes \boldsymbol{I}_{N\times N}.$$

(5.15)
$$\widetilde{\boldsymbol{A}}_n = \widetilde{\boldsymbol{A}}_{n,1} + \widetilde{\boldsymbol{A}}_{n,2},$$

$$\widetilde{\boldsymbol{A}}_{n,1} = \begin{bmatrix} \boldsymbol{A}_0 & 0 & \dots & 0 & \boldsymbol{A}_1 & 0 & \dots & 0 & \boldsymbol{A}_2 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{bmatrix},$$

$$\widetilde{\boldsymbol{A}}_{n,2} = \begin{bmatrix} \widetilde{\boldsymbol{A}}_{n,2,a} & \widetilde{\boldsymbol{A}}_{n,2,b} \\ 0_{n_2N\times n_1N} & \widetilde{\boldsymbol{A}}_{n,2,c} \end{bmatrix},$$

$$\widetilde{\boldsymbol{A}}_{n,2,a} = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & -1 \\ 0 & \vdots & 0 & 1 & -1/2 \end{bmatrix}_{n_1+1\times(n_1+1)} \otimes \boldsymbol{I}_{N\times N},$$

$$\widetilde{\boldsymbol{A}}_{n,2,b} = \begin{bmatrix} 0_{n_1N\times n_2N} \\ [\,-1/2\boldsymbol{I}_{N\times N} \quad 0 \quad \dots \quad 0\,] \end{bmatrix},$$

$$\widetilde{\boldsymbol{A}}_{n,2,c} = \begin{bmatrix} 1/2 & 0 & -1/2 & 0 & \dots \\ 0 & 1/2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & -1/2 \\ 0 & \vdots & 0 & 1/2 & -1/2 \end{bmatrix}_{n_2\times(n_2+1)} \otimes \boldsymbol{I}_{N\times N}.$$

(5.16)
$$\boldsymbol{W\!\!W}_n = \begin{bmatrix} \boldsymbol{W\!\!W}_{n,a} & \boldsymbol{W\!\!W}_{n,b} \\ \boldsymbol{W\!\!W}^{\mathrm{T}}_{n,b} & \boldsymbol{W\!\!W}_{n,c} \end{bmatrix}$$

with

$$\boldsymbol{W}_{n,a} = \frac{r_1}{n_1}\begin{bmatrix} 2/3 & 1/3 & 0 & \dots & 0 \\ 1/3 & 4/3 & 1/3 & \ddots & \vdots \\ 0 & 1/3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 4/3 & 1/3 \\ 0 & \cdot & 0 & 1/3 & 4/3 \end{bmatrix}_{n_1\times n_1} \otimes \boldsymbol{I}_{N\times N},$$

$$\boldsymbol{W}_{n,b} = \begin{bmatrix} \boldsymbol{0}_{(n_1-1)N\times(n_2+1)N} \\ [\, \frac{r_1}{6n_1}\boldsymbol{I}_{N\times N} \quad 0 \quad \dots \quad 0 \quad \frac{r_1}{6n_1}\boldsymbol{A}_1^{\mathrm{T}} \,] \end{bmatrix}_{n_1N\times(n2+1)N},$$

$$\boldsymbol{W}_{n,c} = \boldsymbol{W}_{n,c,1} + \boldsymbol{W}_{n,c,2},$$

$$\boldsymbol{W}_{n,c,1} = \left(\frac{r-r_1}{n_2}\right)\begin{bmatrix} 1/3 & 1/6 & 0 & \dots & 0 \\ 1/6 & 2/3 & 1/6 & \ddots & \vdots \\ 0 & 1/6 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2/3 & 1/6\boldsymbol{A}_2^{\mathrm{T}} \\ 0 & \cdot & 0 & 1/6\boldsymbol{A}_2 & 0 \end{bmatrix}_{(n_2+1)\times(n_2+1)} \otimes \boldsymbol{I}_{N\times N},$$

$$\boldsymbol{W}_{n,c,2} = \begin{bmatrix} \frac{r_1}{6n_1}\boldsymbol{I}_{N\times N} & 0 & \dots & 0 & \frac{r_1}{6n_1}\boldsymbol{A}_1^{\mathrm{T}} \\ 0 & \dots & \dots & \dots & 0 \\ \frac{r_1}{6n_1}\boldsymbol{A}_1 & 0 & \dots & 0 & \boldsymbol{I} + \frac{r_1}{6n_1}\boldsymbol{A}_1\boldsymbol{A}_1^{\mathrm{T}} + \frac{r-r_1}{3n_2}\boldsymbol{A}_2\boldsymbol{A}_2^{\mathrm{T}} \end{bmatrix}.$$

$$\widetilde{\boldsymbol{T}}_{n,2} = \begin{bmatrix} \widetilde{\boldsymbol{T}}_{n,2,a} & \widetilde{\boldsymbol{T}}_{n,2,b} \\ \widetilde{\boldsymbol{T}}_{n,2,c} & \widetilde{\boldsymbol{T}}_{n,2,d} \\ \widetilde{\boldsymbol{T}}_{n,2,e} & \widetilde{\boldsymbol{T}}_{n,2,f} \\ \widetilde{\boldsymbol{T}}_{n,2,g} & \widetilde{\boldsymbol{T}}_{n,2,h} \end{bmatrix}_{(n_1+n_2+1)N\times(n_1+n_2+1)N}, \tag{5.17}$$

$$\widetilde{\boldsymbol{T}}_{n,2,a} = \left(\frac{r_1}{n_1}\right)\left[\begin{bmatrix} 2/3 & 1/3 & 0 & \dots & 0 \\ 1/3 & 4/3 & 1/3 & \ddots & \vdots \\ 0 & 1/3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 4/3 & 1/3 \\ 0 & \cdot & 0 & 1/3 & 4/3 \end{bmatrix}_{n_1\times n_1} \otimes \boldsymbol{I}_{N\times N} \quad \begin{bmatrix} \boldsymbol{0} \\ \frac{1}{3}\boldsymbol{I}_{N\times N} \end{bmatrix}\right]_{n_1N\times(n_1+1)N},$$

$$\widetilde{\boldsymbol{T}}_{n,2,b} = 0_{n_1N\times n_2N},$$

$$\widetilde{\boldsymbol{T}}_{n,2,c} = [\, 0 \quad \dots \quad 0 \quad \frac{r_1}{6n_1}\boldsymbol{I}_{N\times N} \quad \frac{r-r_1}{3n_2}\boldsymbol{I}_{N\times N} + \frac{r_1}{3n_1}\boldsymbol{I}_{N\times N} \,]_{N\times(n_1+1)N},$$

$$\widetilde{\boldsymbol{T}}_{n,2,d} = [\, \frac{r-r_1}{6n_2}\boldsymbol{I}_{N\times N} 0 \quad \dots \quad 0 \,]_{N\times n_2N},$$

$$\widetilde{\boldsymbol{T}}_{n,2,e} = \begin{bmatrix} 0 & \dots & 0 & \frac{r-r_1}{6n_2}\boldsymbol{I}_{N\times N} \\ 0 & \dots & \dots & 0 \end{bmatrix}_{(n_2-1)N\times(n_1+1)N},$$

$$\widetilde{\boldsymbol{T}}_{n,2,f} = \left(\frac{r-r_1}{n_2}\right)\left[\begin{bmatrix} 2/3 & 1/6 & 0 & \dots \\ 1/6 & \ddots & \ddots & 0 \\ 0 & \ddots & 2/3 & 1/6 \\ \dots & 0 & 1/6 & 2/3 \end{bmatrix}_{(n_2-1)\times n_2} \otimes \boldsymbol{I}_{N\times N} \quad \begin{bmatrix} \boldsymbol{0} \\ \frac{1}{6}\boldsymbol{I}_{N\times N} \end{bmatrix}\right]_{(n_2-1)N\times(n_2-1)N},$$

$$\widetilde{\boldsymbol{T}}_{n,2,g} = [\, \boldsymbol{I}_{N\times N} \quad \boldsymbol{0} \quad \tfrac{r_1}{6n_1}\boldsymbol{A}_1 \quad \tfrac{r_1}{3n_1}\boldsymbol{A}_1 \,]_{N\times(n_1+1)N},$$

$$\widetilde{\boldsymbol{T}}_{n,2,h} = [\, \boldsymbol{0} \quad \dots \quad 0 \quad \tfrac{r-r_1}{6n_2}\boldsymbol{A}_2 \quad \tfrac{r-r_1}{3n_2}\boldsymbol{A}_2 \,]_{N\times n_2 N}.$$

(5.18)
$$\widetilde{\mathcal{A}}_n = \widetilde{\mathcal{A}}_{n,1} + \widetilde{\mathcal{A}}_{n,2},$$

$$\widetilde{\mathcal{A}}_{n,1} = \begin{bmatrix} 0 & \dots & \dots & \dots & 0 \\ 2\boldsymbol{I}_{N\times N} & 0 & \dots & 0 & \boldsymbol{A}_0^{\mathrm{T}} \end{bmatrix}_{(n_1+n_2+1)N\times(n_1+n_2+1)N},$$

$$\widetilde{\mathcal{A}}_{n,2} = \begin{bmatrix} \widetilde{\mathcal{A}}_{n,2,a} & \widetilde{\mathcal{A}}_{n,2,b} \\ \widetilde{\mathcal{A}}_{n,2,c} & \widetilde{\mathcal{A}}_{n,2,d} \end{bmatrix}_{(n_1+n_2+1)N\times(n_1+n_2+1)N},$$

$$\widetilde{\mathcal{A}}_{n,2,a} = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 1 \\ 0 & \vdots & 0 & -1 & 0 \end{bmatrix}_{n_1\times n_1} \otimes \boldsymbol{I}_{N\times N},$$

$$\widetilde{\mathcal{A}}_{n,2,b} = \begin{bmatrix} 0 & \dots & \dots & \dots & 0 \\ 1/2\boldsymbol{I}_{N\times N} & 0 & \dots & 0 & 1/2\boldsymbol{A}_1^{\mathrm{T}} \end{bmatrix}_{n_1 N\times(n_2+1)N},$$

$$\widetilde{\mathcal{A}}_{n,2,c} = \begin{bmatrix} 0 & \dots & 0 & -1/2\boldsymbol{I}_{N\times N} \\ \vdots & \vdots & \vdots & 0 \\ 0 & \dots & 0 & -1/2\boldsymbol{A}_1 \end{bmatrix}_{(n_2+1)N\times n_1 N},$$

$$\widetilde{\mathcal{A}}_{n,2,d} = \left[\begin{array}{c|c} \begin{bmatrix} -1/4 & 1/2 & 0 & \dots & 0 \\ -1/2 & 0 & 1/2 & \ddots & \vdots \\ 0 & -1/2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 1/2 \\ 0 & \vdots & 0 & -1/2 & 0 \end{bmatrix}_{n_2\times n_2} \otimes \boldsymbol{I}_{N\times N} & \begin{matrix} 1/4\boldsymbol{A}_1^{\mathrm{T}} \\ 0 \\ \vdots \\ 0 \\ \frac{1}{2}\boldsymbol{A}_2^{\mathrm{T}} \end{matrix} \\ \hline 1/4\boldsymbol{A}_1 \quad 0 \quad \dots \quad 0 \quad -\frac{1}{2}\boldsymbol{A}_2 & \frac{1}{4}\boldsymbol{A}_1\boldsymbol{A}_1^{\mathrm{T}} + 1/2\boldsymbol{A}_2\boldsymbol{A}_2^{\mathrm{T}} \end{array}\right].$$

Remark 5.2. In the case of just one pure delay, matrices in (5.11) are much simpler, due to the fact that vectors v and w are much simpler. Matrices which involve v vectors have been computed in [3]. Here, just to have an idea of such a simplification, matrix $\widetilde{\mathcal{A}}_n$ in (5.18) is reported in the case of one pure delay term.

$$\widetilde{\mathcal{A}}_n = \widetilde{\mathcal{A}}_{n,1} + \widetilde{\mathcal{A}}_{n,2}, \tag{5.19}$$

$$\widetilde{\mathcal{A}}_{n,1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{I}_{N\times N} & \mathbf{0} & \boldsymbol{A}_0^{\mathrm{T}} \end{bmatrix},$$

$$\widetilde{\mathcal{A}}_{n,2} = \left[\begin{array}{c|c} \begin{bmatrix} -1/2 & 1/2 & 0 & \dots & 0 \\ -1/2 & 0 & 1/2 & \ddots & \vdots \\ 0 & -1/2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 1/2 \\ 0 & \vdots & 0 & -1/2 & 0 \end{bmatrix}_{n\times n} \otimes \boldsymbol{I}_{N\times N} & \begin{matrix} \mathbf{0} \\ \frac{1}{2}\boldsymbol{A}_1^{\mathrm{T}} \end{matrix} \\ \hline \mathbf{0} \quad -\frac{1}{2}\boldsymbol{A}_1 & \frac{1}{2}\boldsymbol{A}_1\boldsymbol{A}_1^{\mathrm{T}} \end{array}\right].$$

If there is the distributed delay too, then the following matrix must be added in the right-hand side of (5.19):

$$\widetilde{\mathcal{A}}_{n,3} = \begin{bmatrix} \mathbf{0} & (\boldsymbol{D}_0'^n)^{\mathrm{T}} \\ \vdots & \vdots \\ \mathbf{0} & (\boldsymbol{D}_{n-1}'^n)^{\mathrm{T}} \\ \mathbf{0} & \boldsymbol{A}_1(\boldsymbol{D}_n'^n)^{\mathrm{T}} \end{bmatrix},$$

where

$$\boldsymbol{D}_j'^n = \int_{-r}^0 \boldsymbol{A}_{01}(s)\mathit{spline}_j(s)ds \qquad j = 0, 1, \dots, n.$$

6. Remarks on the infinite horizon case. The methodology here presented for LQG control of hereditary systems over a finite time-horizon can be applied also for LQG control over infinite time-horizon. The basis is the paper [14] in which, under suitable conditions, the convergence of the solution of an approximate Riccati differential equation, evaluated in a sufficiently large time, to the solution of the corresponding infinite dimensional algebraic Riccati equation is proved. The hypotheses required in [14] for such a convergence are satisfied by hereditary systems and by the approximation scheme here presented. Such hypotheses are the Hilbert–Schmidt property of operators $\boldsymbol{Q}$ and $\boldsymbol{F}\boldsymbol{F}^*$, the convergence of the sequence of projection operators involved, and the convergence of the semigroups approximating the semigroup generated by the operator which, in the algebraic Riccati equation, multiplies on the left the unknown Riccati operator. Structural properties are requested in paper [14] of approximate controllability of pairs $(\boldsymbol{A}, \boldsymbol{F})$ and $(\boldsymbol{A}^*, \boldsymbol{Q})$. In that paper the approximate solution of an algebraic Riccati equation is found by exploiting the approximability of the corresponding dynamical Riccati equation and its time convergence toward the steady state, and finding a large enough time-horizon T to approximate the steady-state solution. Such a solving method, which requires only convergence of one approximating semigroup, does not allow for a uniform convergence of the approximate solution toward the actual one (see Theorem 3.2 of [14], relationship among ϵ, T, and n). On the other hand, such a solving method does not require the uniform exponential stability of the approximating semigroups nor the convergence of the adjoint approximate semigroups.

Using the approximate solutions of the Riccati algebraic equations, by using the above paper, the infinite horizon LQG controller can be built. The problem of guaranteeing the convergence of the approximation schemes in infinite horizon case continues to be worthy of attention.

Nevertheless, when the state is fully available, the approximation scheme here presented has the nice property to guarantee convergence also in the infinite horizon case, as stated in the following theorem.

THEOREM 6.1. *Consider system* (2.5)*, with fully available state, that is,*

$$\boldsymbol{y}(t) = \boldsymbol{x}(t) \tag{6.1}$$

and the following cost functional

$$J_I(\boldsymbol{u}) = \lim_{t_f\to\infty} \frac{1}{t_f} \int_0^{t_f} E[(\boldsymbol{Q}\boldsymbol{x}(t), \boldsymbol{x}(t)) + \boldsymbol{u}^{\mathrm{T}}(t)\boldsymbol{u}(t)]dt, \tag{6.2}$$

with $\boldsymbol{Q} : \boldsymbol{M}_2 \mapsto \boldsymbol{M}_2$ *as in* (3.2)*. Let the pair* $(\boldsymbol{A}, \boldsymbol{B})$ *be stabilizable and the pair* $(\boldsymbol{A}, \boldsymbol{Q})$ *be detectable. Let*

$$\boldsymbol{u}_n(t) = -\boldsymbol{B}^\star \boldsymbol{R}_n(T)\boldsymbol{\Pi}'_n \boldsymbol{x}_n(t), \tag{6.3}$$

where $\boldsymbol{R}_n(T)$ *is the approximate solution of the algebraic Riccati equation for control*

$$\boldsymbol{A}^\star\boldsymbol{R} + \boldsymbol{R}\boldsymbol{A} - \boldsymbol{R}\boldsymbol{B}\boldsymbol{B}^\star\boldsymbol{R} + \boldsymbol{Q} = 0 \tag{6.4}$$

obtained [14] *by evaluating the approximate dynamic Riccati equation* (3.16) *in a suitable time* T*, and* $\boldsymbol{x}_n(t)$ *is the corresponding evolving state. Let* $\boldsymbol{x}(t)$ *be the state evolving when the optimal infinite horizon LQG control law is applied to the system. Then, for every* $\epsilon > 0$*, there exists a* T_ϵ*, such that for every* $T > T_\epsilon$ *there exists an* n_T*, such that for every* $n > n_T$ *the semigroup which governs the closed loop system, that is, the one generated by* $\boldsymbol{A} - \boldsymbol{B}\boldsymbol{B}^\star\boldsymbol{R}_n(T)\boldsymbol{\Pi}'_n$*, is exponentially stable and, moreover,*

$$E\|\boldsymbol{x}_n(t) - \boldsymbol{x}(t)\| < \epsilon \qquad \forall t \in [0, \infty). \tag{6.5}$$

Proof. First let us prove that $E\|x(t)\|$ is uniformly bounded. Let $S(t)$ be the semigroup generated by the optimal closed loop infinitesimal generator $\boldsymbol{A} - \boldsymbol{B}\boldsymbol{B}^\star\boldsymbol{R}$, with $\boldsymbol{R}$ the solution of the algebraic Riccati control equation. There exist positive constants M and σ such that $\|S(t)\| \le Me^{-\sigma t}$. It is

$$\begin{aligned} E\|\boldsymbol{x}(t)\| &\le E\|S(t)\boldsymbol{x}(0)\| + E\|\int_0^t S(t-\tau)\boldsymbol{F}\boldsymbol{\omega}(\tau)d\tau\| \\ &\le Me^{-\sigma t}\sqrt{E(\|\boldsymbol{x}(0)\|)^2} + \left(\int_0^t M^2 e^{-2\sigma(t-\tau)}\|\boldsymbol{F}\|^2_{H.S.}\right)^{\frac{1}{2}} \\ &\le M\sqrt{\mathrm{Tr}(\boldsymbol{P}_0)} + \frac{M}{\sqrt{2\sigma}}\|\boldsymbol{F}\|_{H.S.} \end{aligned} \tag{6.6}$$

Consider now the equation

$$\dot{\xi}_{n,T}(t) = (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{B}^\star \boldsymbol{R}_n(T)\boldsymbol{\Pi}_n')\xi_{n,T}(t). \tag{6.7}$$

It is

$$\xi_{n,T}(t) = S(t)\xi_{n,T}(0) + \int_0^t S(t-\tau)\boldsymbol{B}\boldsymbol{B}^\star(\boldsymbol{R} - \boldsymbol{R}_n(T)\boldsymbol{\Pi}_n')\xi_{n,T}(\tau)d\tau \tag{6.8}$$

by which it follows that

$$\|\xi_{n,T}(t)\| \le Me^{-\sigma t}\|\xi_{n,T}(0)\| \; + \int_0^t Me^{-\sigma(t-\tau)}\|\boldsymbol{B}\boldsymbol{B}^\star(\boldsymbol{R} - \boldsymbol{R}_n(T)\boldsymbol{\Pi}_n')\|_{H.S.}\xi_{n,T}(\tau)d\tau \tag{6.9}$$

and by the Gronwall inequality

$$\|\xi_{n,T}(t)\| \le Me^{(-\sigma + M\|\boldsymbol{B}\boldsymbol{B}^\star(\boldsymbol{R} - \boldsymbol{R}_n(T)\boldsymbol{\Pi}_n')\|_{H.S.}t)}\xi_{n,T}(0). \tag{6.10}$$

Now let $\epsilon > 0$. By Theorem 3.2 in [14] and by the inequality

$$\|\boldsymbol{R} - \boldsymbol{R}_n(T)\boldsymbol{\Pi}_n'\|_{H.S.} \le \|\boldsymbol{\Pi}_n'\boldsymbol{R}\boldsymbol{\Pi}_n' - \boldsymbol{R}_n(T)\boldsymbol{\Pi}_n'\|_{H.S.} + \|\boldsymbol{R} - \boldsymbol{\Pi}_n'\boldsymbol{R}\boldsymbol{\Pi}_n'\|_{H.S.}, \tag{6.11}$$

it follows that there exists T_ϵ such that for every $T > T_\epsilon$ there exists n_T such that for every $n > n_T$

$$\|\boldsymbol{R} - \boldsymbol{R}_n(T)\boldsymbol{\Pi}_n'\|_{H.S.} < \min\left\{\frac{\sigma}{2M\|\boldsymbol{B}\boldsymbol{B}^\star\|}, \frac{\epsilon\sigma}{2M\|\boldsymbol{B}\boldsymbol{B}^\star\|\sup_{\tau\in[-r,\infty)} E\|\boldsymbol{x}(\tau)\|}\right\} \tag{6.12}$$

and so

$$\|\xi_{n,T}(t)\| \le Me^{\frac{-\sigma}{2}t}\xi_{n,T}(0), \tag{6.13}$$

which implies the exponential stability of the closed loop semigroup.

As far as the second part of the thesis is concerned, let

$$e_n(t) = \boldsymbol{x}(t) - \boldsymbol{x}_n(t).$$

It is

$$\dot{e}_n(t) = (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{B}^\star\boldsymbol{R}_n(T)\boldsymbol{\Pi}_n')e_n(t) + \boldsymbol{B}\boldsymbol{B}^\star(\boldsymbol{R}_n(T)\boldsymbol{\Pi}_n' - \boldsymbol{R})\boldsymbol{x}(t) \tag{6.14}$$

by which, taking into account (6.12),

$$\begin{aligned} E\|e_n(t)\| \le \int_0^t & Me^{(-\sigma + M\|\boldsymbol{B}\boldsymbol{B}^\star(\boldsymbol{R} - \boldsymbol{R}_n(T)\boldsymbol{\Pi}_n')\|_{H.S.})(t-\tau)} \\ & \cdot\|\boldsymbol{B}\boldsymbol{B}^\star(\boldsymbol{R} - \boldsymbol{R}_n(T)\boldsymbol{\Pi}_n')\|_{H.S.}E\|\boldsymbol{x}(\tau)\|d\tau < \epsilon. \qquad \square \end{aligned} \tag{6.15}$$

7. Examples. Simulations reported in this section have been performed by MATLAB on a PC using the 3rd order Runge–Kutta integration algorithm.[1]

[1]Simulation programs are available upon request.

Example 1. Consider the following unstable hereditary system:

$$\begin{aligned}\frac{d^2z(t)}{dt^2} &= \frac{dz(t)}{dt} + z(t) + \frac{dz(t-r_1)}{dt} + z(t-r_1)\\ &\quad + \frac{dz(t-r_2)}{dt} + z(t-r_2) + u(t) + \omega_1(t),\\ y(t) &= z(t) + \omega_2(t),\end{aligned} \tag{7.1}$$

where $z(t), u(t), y(t) \in \mathbb{R}$, $\omega_1(t), \omega_2(t) \in \mathbb{R}$ are independent white Gaussian standard noises.

By denoting $Z(t) = \left[\begin{smallmatrix} z(t) \\ \frac{dz(t)}{dt} \end{smallmatrix}\right]$, and $\omega(t) = \left[\begin{smallmatrix} \omega_1(t) \\ \omega_2(t) \end{smallmatrix}\right]$, the system (7.1) can be rewritten as follows:

$$\begin{aligned}\dot{Z}(t) &= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} Z(t) + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} Z(t-r_1)\\ &\quad + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} Z(t-r_2) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \omega(t)\\ y(t) &= [\,1 \quad 0\,]\, Z(t) + [\,0 \quad 1\,]\, \omega(t).\end{aligned} \tag{7.2}$$

The weight matrix in the functional (3.1) has been chosen as

$$\boldsymbol{Q}_0 = \begin{bmatrix} 1000 & 0 \\ 0 & 0 \end{bmatrix}. \tag{7.3}$$

The time t_f has been chosen equal to 10, and the delays have been chosen as $r_1 = 1.2$ and $r_2 = 2.5$. The initial value of $Z(\vartheta), \vartheta \in [-r_2, 0]$, has been chosen as follows:

$$Z(\vartheta) = \begin{bmatrix} e^{-\vartheta} \\ \cos(10\vartheta) \end{bmatrix},$$

while the initial estimate $\hat{Z}(\vartheta)$ has been set to 0 in the same interval. The covariance operator $\boldsymbol{P}_0$ of the initial state in M_2 has been chosen as follows:

$$\boldsymbol{P}_0\boldsymbol{x} = (\boldsymbol{x}, \phi)\phi,$$

where $\boldsymbol{x}, \phi \in M_2$, $\phi = \left[\begin{smallmatrix} \phi_0 \\ \phi_1 \end{smallmatrix}\right]$, $\phi_0 = \phi_1(\vartheta) = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right], \vartheta \in [-r_2, 0]$.

The integration step has been chosen equal to 0.0025.

Figures 2–5 report the first component of the actual $Z(t)$ evolving when the approximated input is applied to the system and of the estimated $\hat{Z}(t)$ for different values of n_1 and n_2. Figures 6–9 report the second component of $Z(t)$ and $\hat{Z}(t)$. In Figures 10–13 the first component and the second one of $Z(t)$ and $\hat{Z}(t)$, the approximated control input and the noisy output, are reported for $n_1 = n_2 = 6$.

Example 2. Consider now the well-known National Transonic Facility [4, 27, 40], the liquid nitrogen wind tunnel at NASA Langley Research Center in Hampton, VA. Here only one of the state variables is measured, the guide vane angle, while no measurement of the mach number nor of the guide vane angle derivative is available. Moreover, we suppose an additive Gaussian noise corrupts the dynamics of the system and the above measure. A simplified model of such a system is given by (see [4] for the

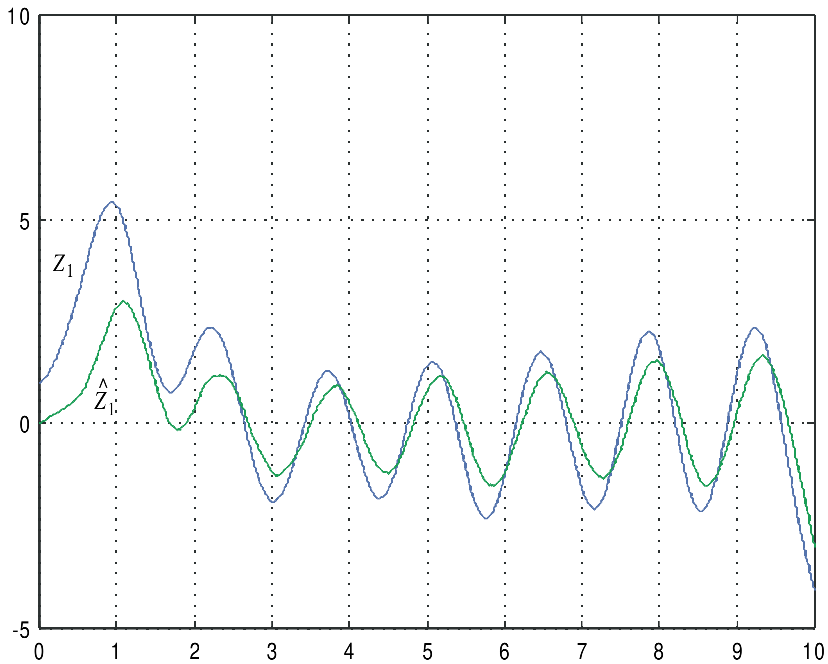

FIG. 2. *The case of* $n_1 = n_2 = 2$*: true and estimated* $Z_1(t)$*.*

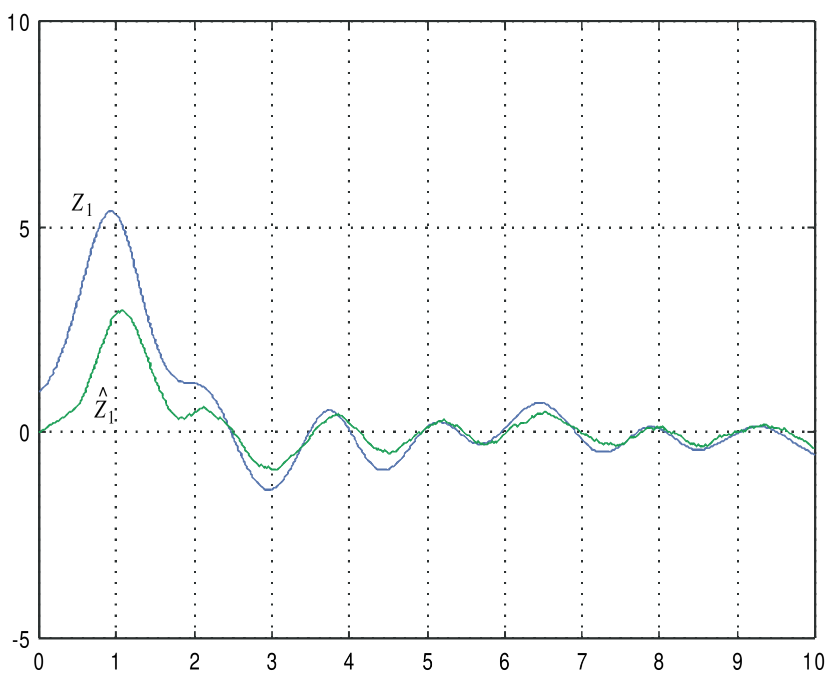

FIG. 3. *The case of* $n_1 = 3$, $n_2 = 2$*: true and estimated* $Z_1(t)$*.*

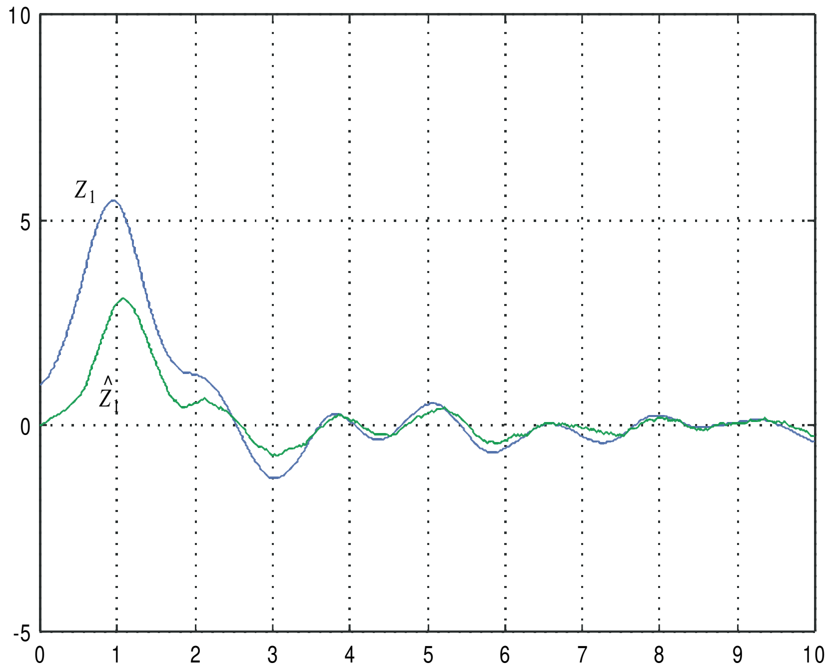

FIG. 4. *The case of* $n_1 = 3$, $n_2 = 3$: *true and estimated* $Z_1(t)$.

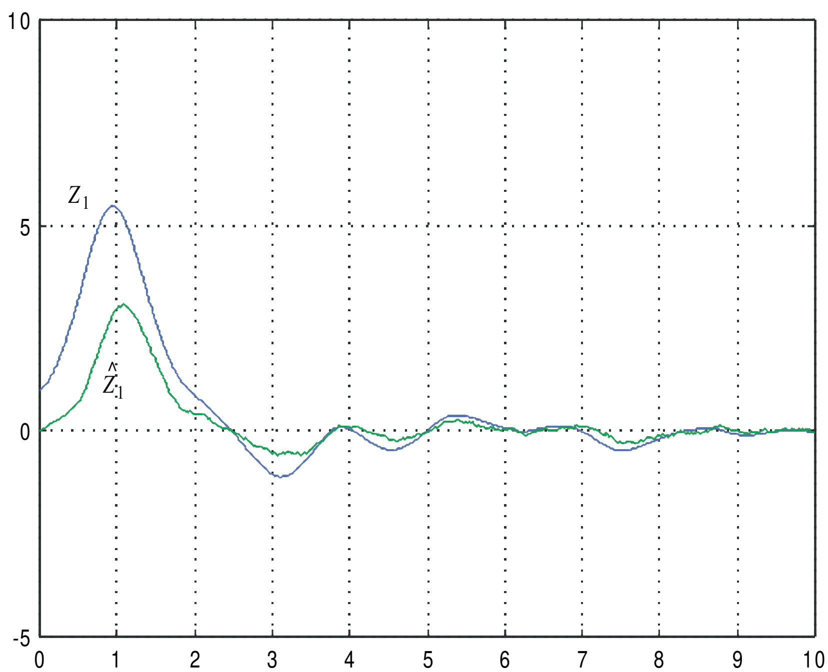

FIG. 5. *The case of* $n_1 = 4$, $n_2 = 3$: *true and estimated* $Z_1(t)$.

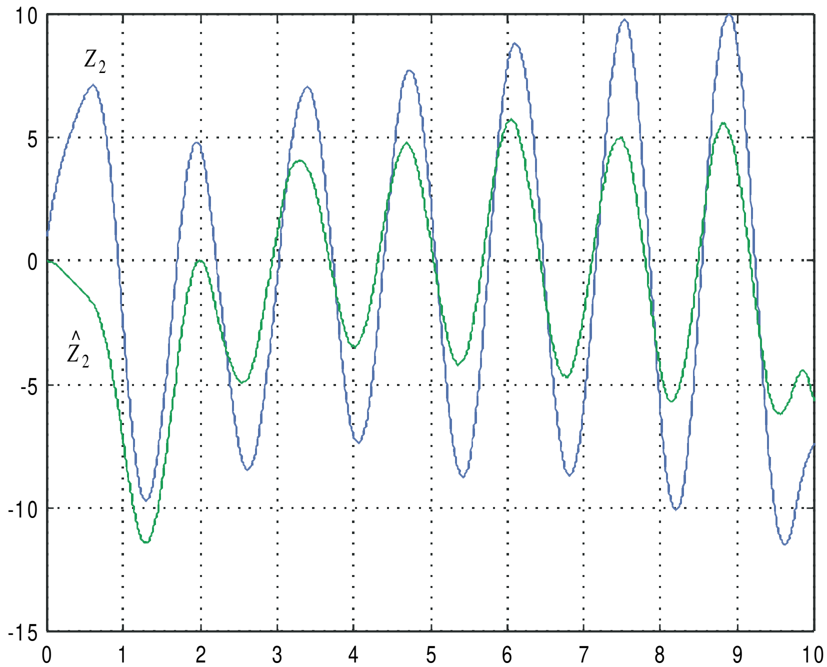

FIG. 6. *The case of* $n_1 = n_2 = 2$*: true and estimated* $Z_2(t)$*.*

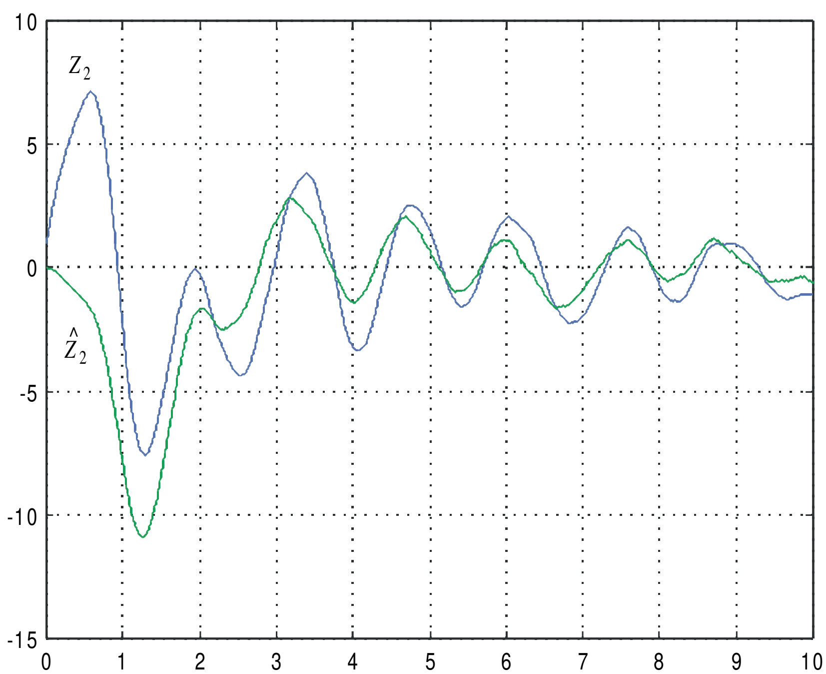

FIG. 7. *The case of* $n_1 = 3$, $n_2 = 2$*: true and estimated* $Z_2(t)$*.*

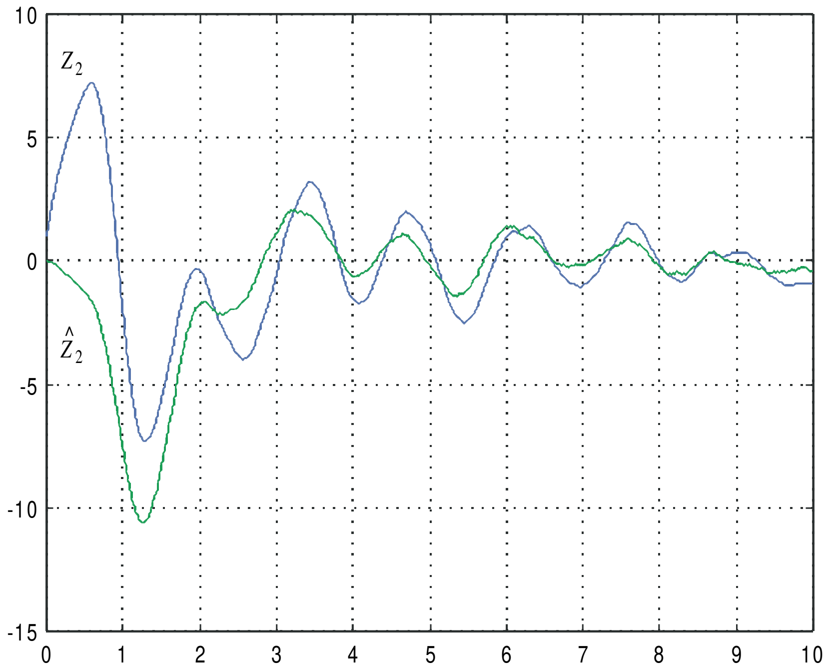

FIG. 8. *The case of* $n_1 = 3$, $n_2 = 3$: *true and estimated* $Z_2(t)$.

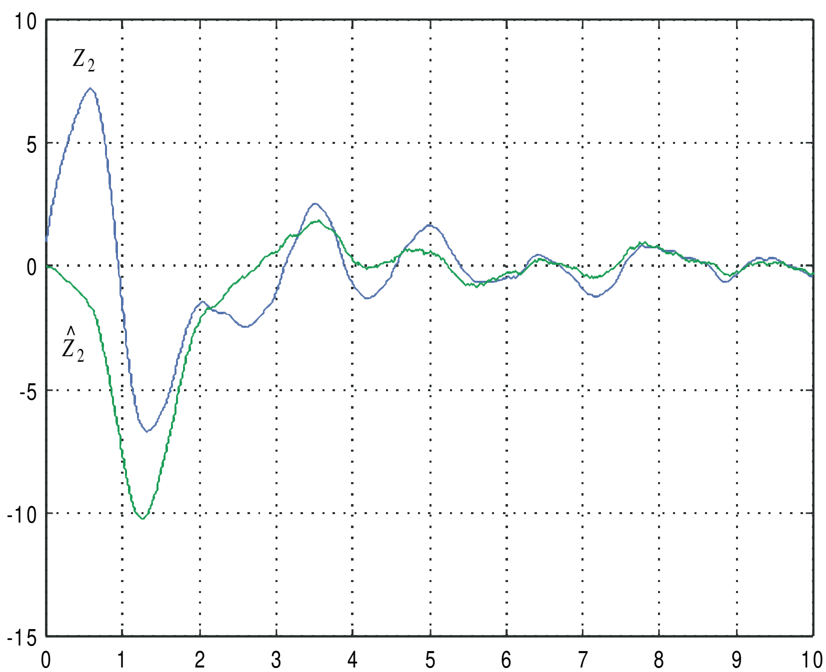

FIG. 9. *The case of* $n_1 = 4$, $n_2 = 3$: *true and estimated* $Z_2(t)$.

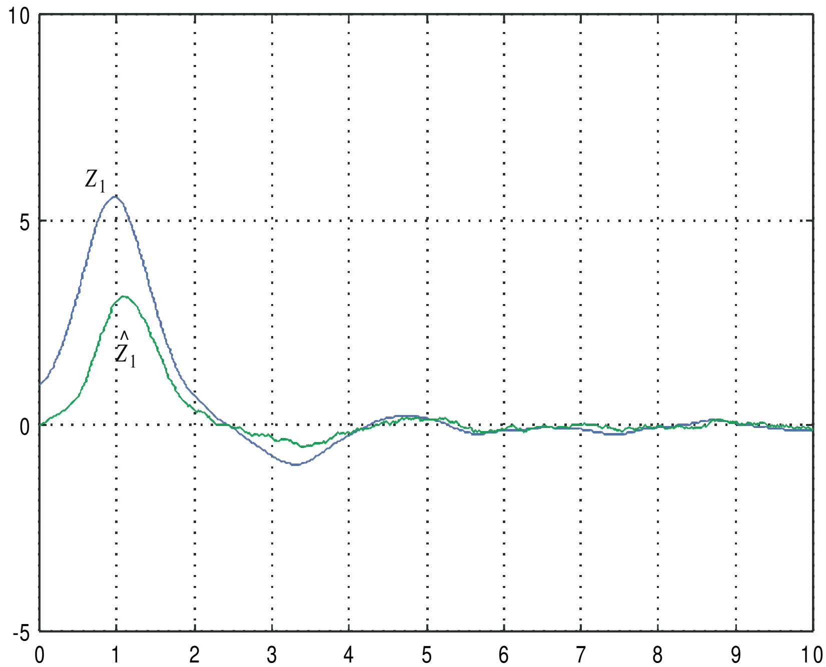

FIG. 10. *The case of* $n_1 = n_2 = 6$: *true and estimated* $Z_1(t)$.

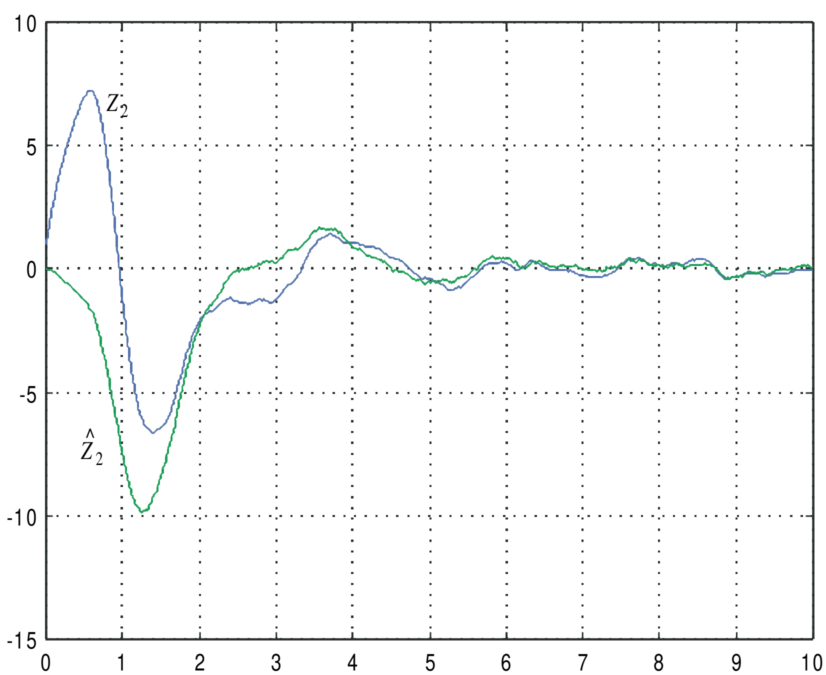

FIG. 11. *The case of* $n_1 = n_2 = 6$: *true and estimated* $Z_2(t)$.

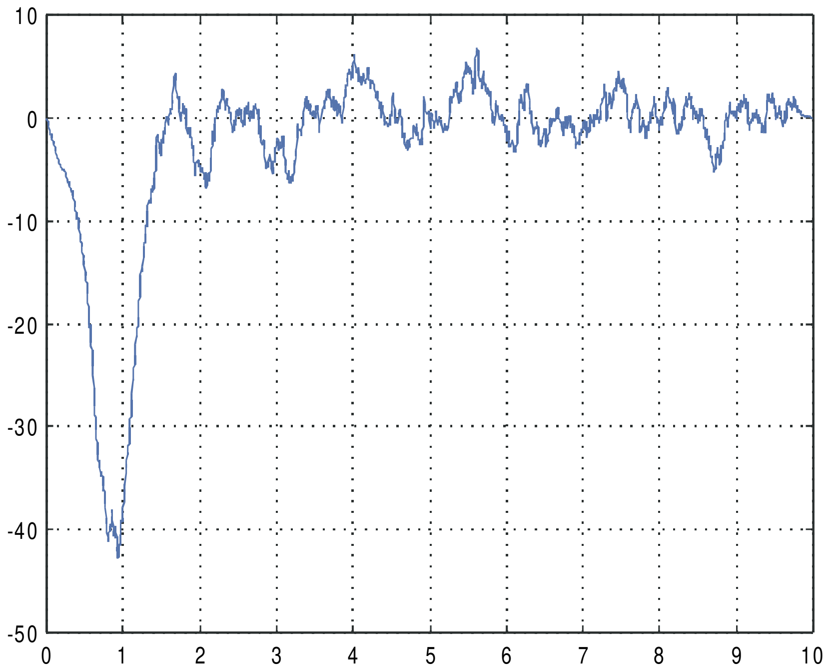

Fig. 12. *The case of* $n_1 = n_2 = 6$: *the input* $u(t)$.

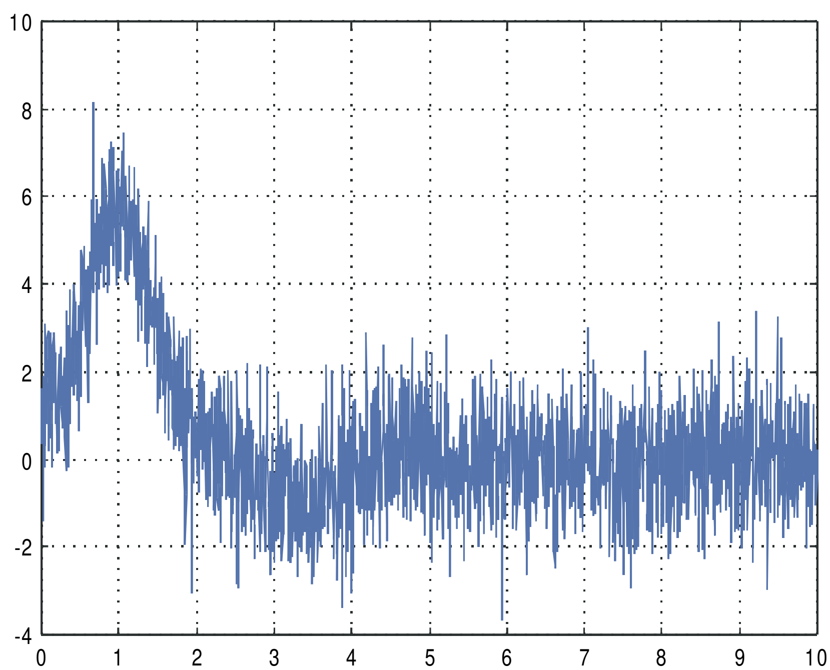

Fig. 13. *The case of* $n_1 = n_2 = 6$:: *the noisy output* $y(t)$.

deterministic model)

$$\dot{z}(t) = \begin{bmatrix} -a & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -\bar{\omega}^2 & -2\xi\bar{\omega} \end{bmatrix} z(t) + \begin{bmatrix} 0 & ka & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} z(t-0.33) \tag{7.4}$$

$$+ \begin{bmatrix} 0 \\ 0 \\ -\bar{\omega}^2 \end{bmatrix} \boldsymbol{u}(t) + F_0\omega_1(t)$$

$$\boldsymbol{y}(t) = [\,0 \quad 1 \quad 0)\,]\, z(t) + G_0\omega_2(t)$$

with $(1/a) = 1.964$, $k = -0.0117$, $\xi = 0.8$, $\bar{\omega} = 6.0$, and $\omega_1(t), \omega_2(t) \in \mathbb{R}$ independent white Gaussian standard noises. As in the LQ problem developed in [4, 27, 40], the matrix $\boldsymbol{Q}_0$ in the functional (3.1) has been chosen as follows:

$$\boldsymbol{Q}_0 = \begin{bmatrix} 10000 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{7.5}$$

In simulations we have supposed to know exactly the initial state

$$z(\tau) = \begin{bmatrix} -0.1 \\ 8.547 \\ 0 \end{bmatrix}, \qquad \tau \in [-0.33, 0],$$

and we have used

$$\boldsymbol{F}_0 = \begin{bmatrix} 0 \\ 0 \\ 10 \end{bmatrix}, \qquad \boldsymbol{G} = 1.$$

Figures 14–17 show the three components of the state and the input for $n = 2$. Computed values of the functional

$$J_{10} = \int_0^{10} [\boldsymbol{z}^{\mathrm{T}}(t)\boldsymbol{Q}_0\boldsymbol{z}(t) + \boldsymbol{u}^2(t)]dt \tag{7.6}$$

for different n and the same noise realization are reported in Table 1. The integration step has been chosen equal to $dT = 0.001$, the integral J_{10} has been computed as $dT\sum_{k=0}^{10/dT} \boldsymbol{z}^{\mathrm{T}}(kdT)\boldsymbol{Q}_0\boldsymbol{z}(kdT) + u^2(kdT)$.

We have considered also the infinite horizon LQ problem: this means that we have considered only the Riccati equation for control and evaluated the dynamic approximated Riccati operator in a sufficiently large time. We have stopped integration when the norm of the difference between the Riccati matrix operators evaluated in time kdT and $(k+1)dT$ was less than 10^{-10}.

Tables 2–5 report the values of matrices Π_0^n and of matrices of functions $\Pi_1^n(\vartheta)$ [4, 27, 40] of the approximated, not yet implementable, LQ control law

$$u_n(t) = \Pi_0^n z(t) + \int_{-r}^0 \Pi_1^n(\theta) z(t+\theta)d\theta. \tag{7.7}$$

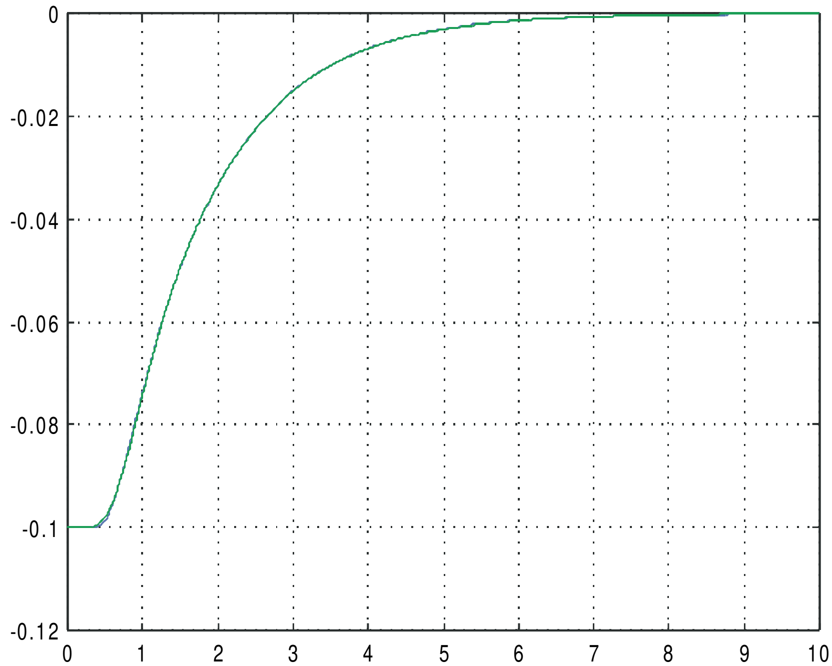

FIG. 14. *Finite horizon LQG for the wind tunnel. The case of* $n = 2$*: true and estimated* $z_1(t)$ *(almost coincident).*

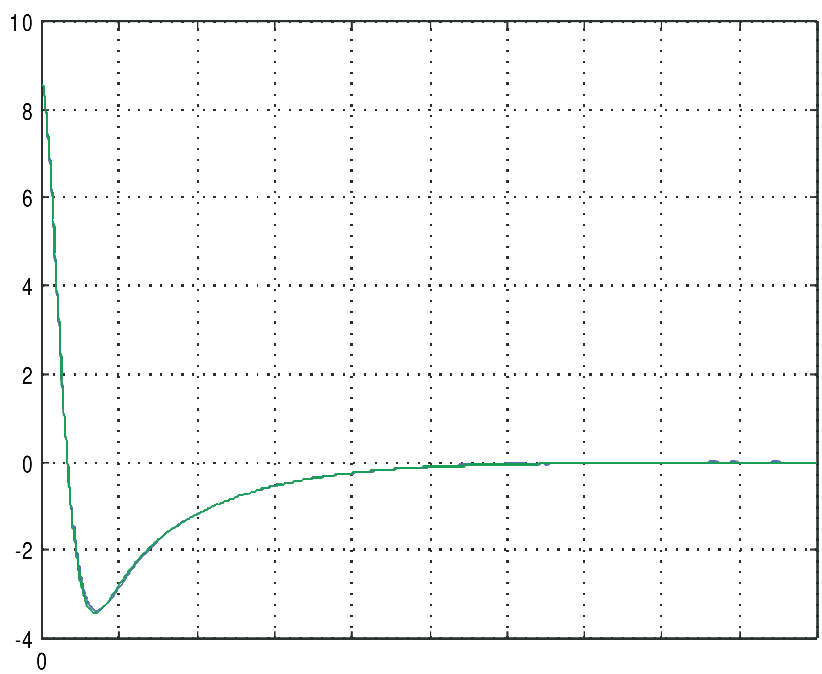

FIG. 15. *Finite horizon LQG for the wind tunnel. The case of* $n = 2$*: true and estimated* $z_2(t)$ *(almost coincident).*

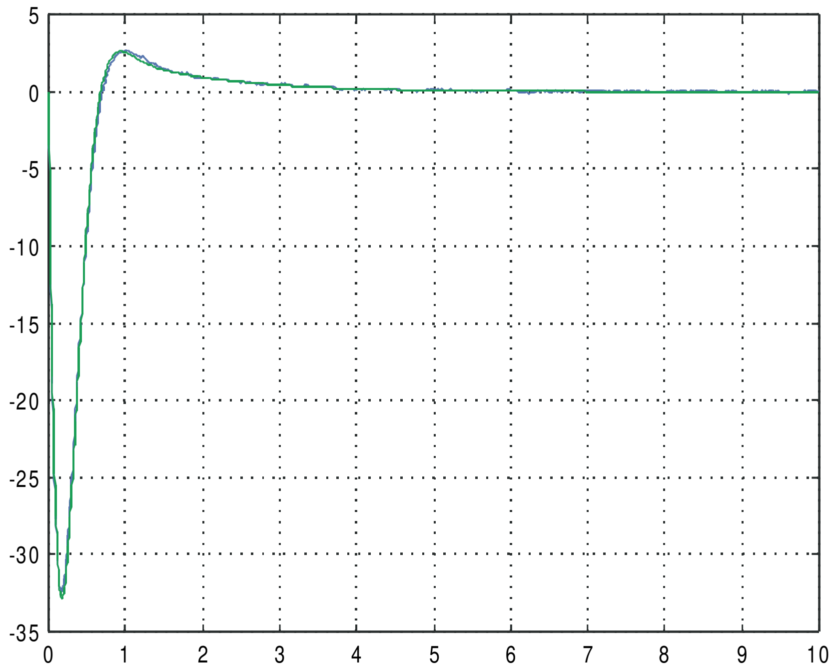

Fig. 16. *Finite horizon LQG for the wind tunnel. The case of $n = 2$: true and estimated $z_3(t)$ (almost coincident).*

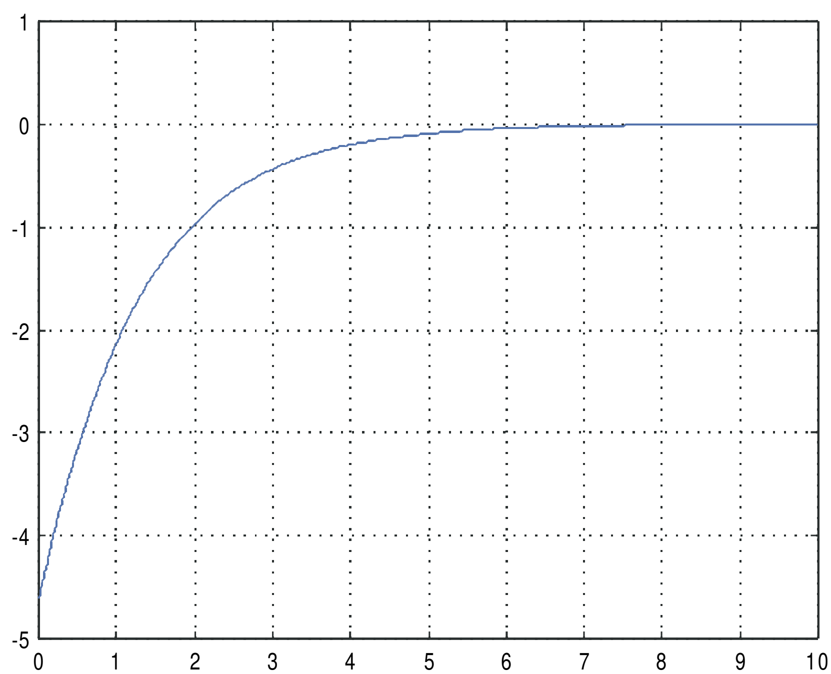

Fig. 17. *Finite horizon LQG for the wind tunnel. The case of $n = 2$: the input $u(t)$.*

Table 1
Values of J_{10} computed for different values of n.

n	J_{10}
2	136.41324
4	136.41311
8	136.41310

Table 2
Values of matrix Π_0^n for different values of n.

$n=2$	$\begin{bmatrix} 8676.5662 & -9.8145 & -0.9479 \\ -9.8145 & 0.0182 & 0.0018 \\ -0.9479 & 0.0018 & 0.0002 \end{bmatrix}$
$n=4$	$\begin{bmatrix} 8676.9112 & -9.8149 & -0.9477 \\ -9.8149 & 0.0184 & 0.0018 \\ -0.9477 & 0.0018 & 0.0002 \end{bmatrix}$
$n=8$	$\begin{bmatrix} 8676.9959 & -9.8150 & -0.9477 \\ -9.8150 & 0.0185 & 0.0018 \\ -0.9477 & 0.0018 & 0.0002 \end{bmatrix}$
$n=16$	$\begin{bmatrix} 8677.0170 & -9.8150 & -0.9477 \\ -9.8150 & 0.0185 & 0.0018 \\ -0.9477 & 0.0018 & 0.0002 \end{bmatrix}$

Table 3
Values of $\Pi_1^n(1,2)$ for different values of n.

j	$\Pi_1^2(-jr/16)$	$\Pi_1^4(-jr/16)$	$\Pi_1^8(-jr/16)$	$\Pi_1^{16}(-jr/16)$
0	−41.3798	−41.3931	−41.3962	−41.3970
1	—	—	—	−42.0024
2	—	—	−42.6103	−42.6140
3	—	—	—	−43.2288
4	—	−43.8343	−43.8491	−43.8499
5	—	—	—	−44.4742
6	—	—	−45.1014	−45.1051
7	—	—	—	−45.7393
8	−46.3182	−46.3773	−46.3796	−46.3802
9	—	—	—	−47.0246
10	—	—	−47.6721	−47.6757
11	—	—	—	−48.3306
12	—	−48.9777	−48.9920	−48.9923
13	—	—	—	−49.6579
14	—	—	−50.3271	−50.3306
15	—	—	—	−51.0071
16	−51.6883	−51.6904	−51.6909	−51.6910

TABLE 4
Values of $\Pi_1^n(2,2)$ *for different values of* n.

j	$\Pi_1^2(-jr/16)$	$\Pi_1^4(-jr/16)$	$\Pi_1^8(-jr/16)$	$\Pi_1^{16}(-jr/16)$
0	0.0684	0.0690	0.0691	0.0692
1	—	—	—	0.0685
2	—	—	0.0678	0.0677
3	—	—	—	0.0670
4	—	0.0665	0.0663	0.0663
5	—	—	—	0.0656
6	—	—	0.0650	0.0649
7	—	—	—	0.0643
8	0.0646	0.0635	0.0636	0.0636
9	—	—	—	0.0629
10	—	—	0.0623	0.0623
11	—	—	—	0.0616
12	—	0.0613	0.0610	0.0610
13	—	—	—	0.0604
14	—	—	0.0598	0.0597
15	—	—	—	0.0591
16	0.0585	0.0585	0.0585	0.0585

TABLE 5
Values of $\Pi_1^n(3,2)$ *for different values of* n.

j	$\Pi_1^2(-jr/16)$	$\Pi_1^4(-jr/16)$	$\Pi_1^8(-jr/16)$	$\Pi_1^{16}(-jr/16)$
0	0.0067	0.0067	0.0067	0.0067
1	—	—	—	0.0066
2	—	—	0.0065	0.0065
3	—	—	—	0.0065
4	—	0.0064	0.0064	0.0064
5	—	—	—	0.0063
6	—	—	0.0063	0.0063
7	—	—	—	0.0062
8	0.0062	0.0061	0.0061	0.0061
9	—	—	—	0.0061
10	—	—	0.0060	0.0060
11	—	—	—	0.0060
12	—	0.0059	0.0059	0.0059
13	—	—	—	0.0058
14	—	—	0.0058	0.0058
15	—	—	—	0.0057
16	0.0056	0.0056	0.0056	0.0056

In Tables 3–5 the values of the second column, the only one not zero, of matrices Π_1^n of continuous functions are reported, just in instants $-jr/n$, $j = 0, 1, \dots, n$ (between such points the continuous function in consideration is a one degree polynomial).

In the wind tunnel example, and in all other examples we have simulated, no oscillations appear for $\Pi_1^n(\vartheta)$, which was an important problem arising while consid-

TABLE 6
Values of Π_0^8 for different approximation schemes.

[2, 24]	$\begin{bmatrix} 8671.3161 & -9.8336 & -0.9500 \\ -9.8336 & 0.0179 & 0.0018 \\ -0.9500 & 0.0018 & 0.0002 \end{bmatrix}$
[4]	$\begin{bmatrix} 8676.9829 & -9.8154 & -0.9477 \\ -9.8154 & 0.0185 & 0.0019 \\ -0.9477 & 0.0019 & 0.0002 \end{bmatrix}$
[27]	$\begin{bmatrix} 8677.02698 & -9.81505 & -0.94768 \\ -9.81505 & 0.01851 & 0.00186 \\ -0.94768 & 0.00186 & 0.00019 \end{bmatrix}$
[40]	$\begin{bmatrix} 8677.02502 & -9.81503 & -0.94768 \\ -9.81503 & 0.01851 & 0.00186 \\ -0.94768 & 0.00186 & 0.00019 \end{bmatrix}$
[this paper]	$\begin{bmatrix} 8676.99592 & -9.81502 & -0.94769 \\ -9.81502 & 0.01850 & 0.00186 \\ -0.94769 & 0.00186 & 0.00019 \end{bmatrix}$
Π_0	$\begin{bmatrix} 8677.02405 & -9.81505 & -0.94768 \\ -9.81505 & 0.01851 & 0.00186 \\ -0.94768 & 0.00186 & 0.00019 \end{bmatrix}$

TABLE 7
Values of $J(u_n)$ for different schemes and $n = 8$.

[2, 24]	[4]	[27]	[40]	[This paper]	Theor. value [27]
136.7361	136.7354	136.40094	136.4490	136.4131	136.40490

ering the approximation scheme [4]. In that approximation scheme Π_1^n is increasingly oscillatory with increasing n, while in our scheme, as in [27, 40], each function in Π_1^n is monotone. This property becomes very important if one wants to implement the approximating feedback law in a real system [27].

In Table 6 the approximations of order $n = 8$ of matrix Π_0, denoted Π_0^8, are reported, computed with approximation schemes in [2, 24, 4, 27, 40] and with the method presented in this paper. Also the exact optimal Π_0 is reported, as computed in [27].

Table 7 reports the values of the functional

$$\int_0^\infty [\boldsymbol{z}^{\mathrm{T}}(t)\boldsymbol{Q}_0\boldsymbol{z}(t) + \boldsymbol{u}^2(t)]dt \tag{7.8}$$

computed using the same approximation schemes for $n = 8$.

The value computed with the method proposed in this paper is obtained by numerical integration of (7.6). The value computed with $t_f = 20$ is quite the same (for $n = 2$ it is $J_{20} = 136.4133$).

TABLE 8
Numerical values of J_{10} computed for different values of n.

n	J_{10}
2	136.41325
4	136.41312
8	136.41311

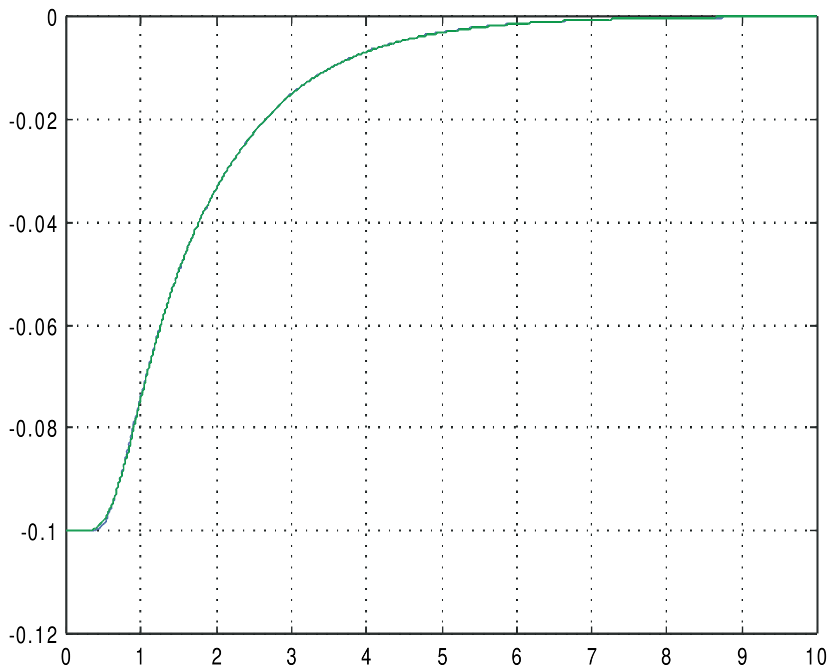

FIG. 18. *Infinite horizon LQG for the wind tunnel. The case of $n = 2$: true and estimated $z_1(t)$ (almost coincident).*

To conclude, the infinite horizon LQG control is considered: the solutions of the approximate algebraic Riccati equations for control and filtering are used in the control scheme. The resulting controller is a dynamic finite dimensional stationary system driven by the noisy output. The values of the index $\int_0^{10}[\boldsymbol{z}^{\mathrm{T}}(t)\boldsymbol{Q}_0\boldsymbol{z}(t)+\boldsymbol{u}^2(t)]dt$, computed for different n and for the same noise realization, are reported in Table 8.

All approximation schemes give, within numerical errors, practically the same value of the cost functional. It is indeed remarkable that the proposed approximation scheme is able to reach such value of the functional starting from noisy output measurements and not, as the other schemes do, starting from noiseless full state information (in a delay interval).

In Figures 18–20 the three components of the state are reported in the infinite horizon case, for $n = 2$. The plots of the input and of the output are reported in Figures 21 and 22.

Comparison with the methods presented in [25, 26] cannot be reported because such papers do not contain numerical results.

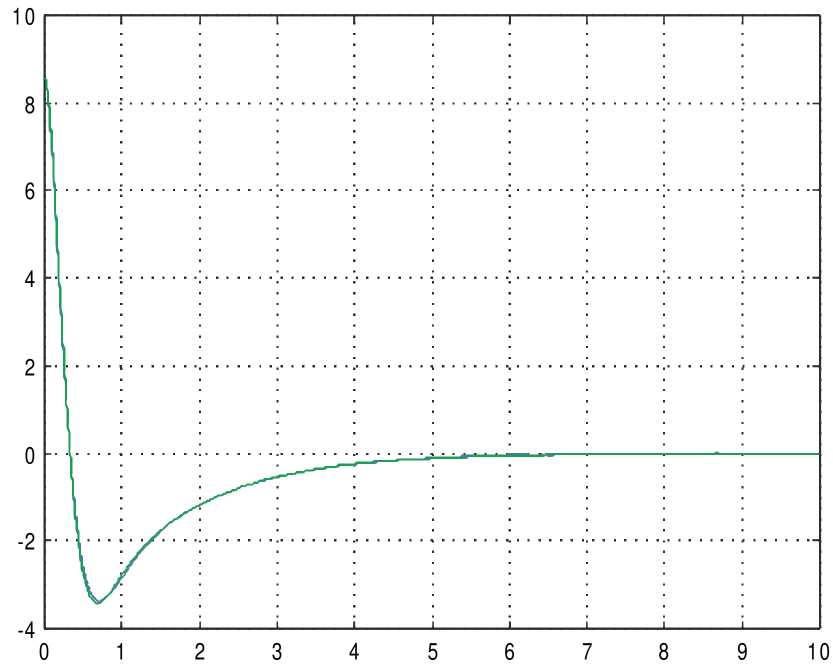

FIG. 19. *Infinite horizon LQG for the wind tunnel. The case of* $n = 2$*: true and estimated* $z_2(t)$ *(almost coincident).*

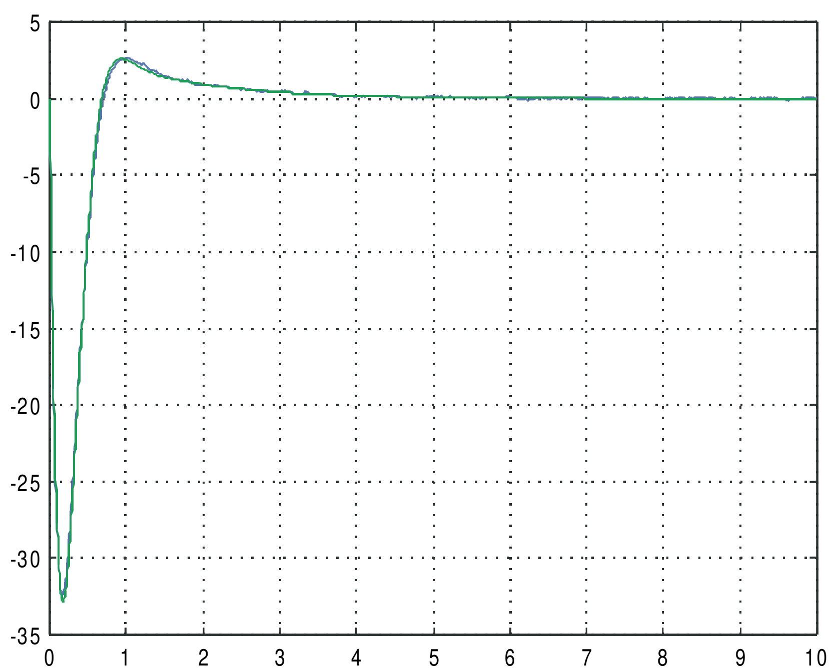

FIG. 20. *Infinite horizon LQG for the wind tunnel. The case of* $n = 2$*: true and estimated* $z_3(t)$ *(almost coincident).*

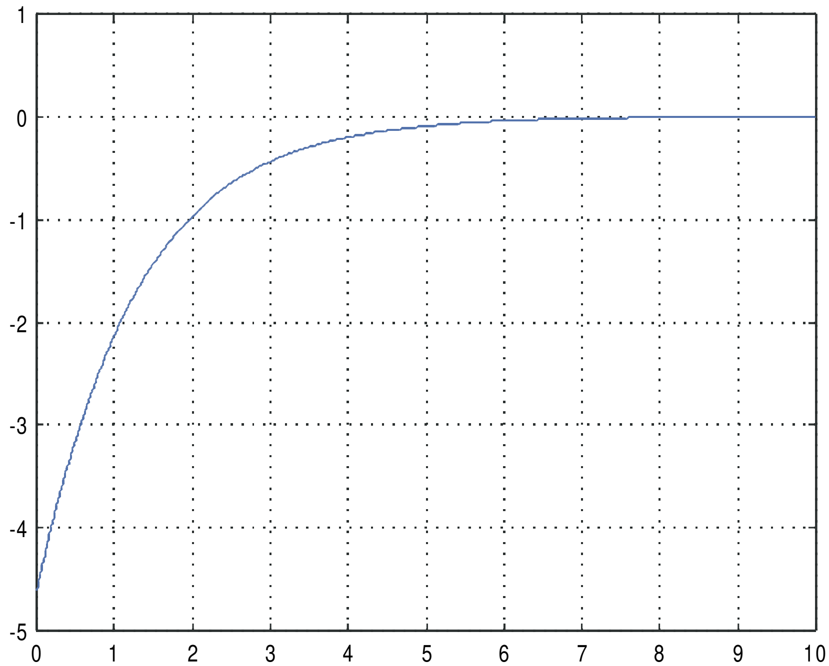

FIG. 21. *Infinite horizon LQG for the wind tunnel. The case of $n = 2$: the input $u(t)$.*

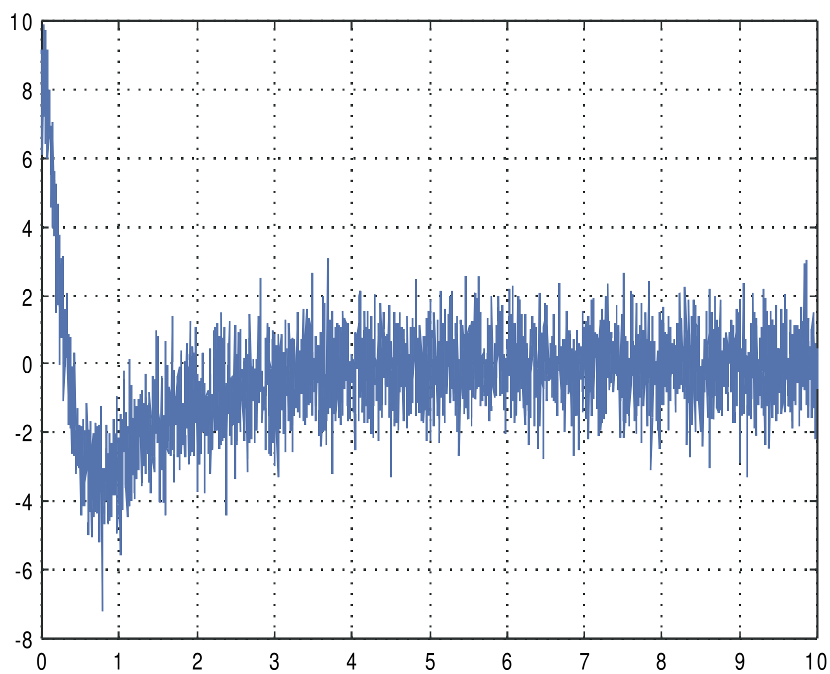

FIG. 22. *Infinite horizon LQG for the wind tunnel. The case of $n = 2$: the noisy output $y(t)$.*

8. Conclusions. In this paper a new spline approximation scheme has been developed for the finite horizon LQG control of general hereditary systems. The approximated LQG controller is a finite dimensional linear system driven by the noisy output. It has been proved that the approximated implementable feedback input and the corresponding state converge to the optimal ones. The approximation scheme makes use of first order splines, introduced by [3] for hereditary systems, suitably adapted to the LQG problem which involves three differential equations, that is, the filter equation and the two Riccati equations required for the computation of the optimal stochastic control. Generally in the literature one of these equations is considered, that is the Riccati equation for the deterministic state feedback optimal control [4, 24, 26, 27, 40]. A methodology with two approximating subspaces has been necessary to apply such spline functions to obtain convergence of the overall LQG problem.

The main feature of the proposed scheme is from a numerical point of view. Indeed our proposal of an implementable LQG controller gives practically the same results of the well-known LQ controller, with a complete knowledge of the infinite dimensional state in a deterministic setting, with reference to an important widely studied case as the NASA National Transonic Facility. The choice of spline environment instead of averaging one is motivated in [4], where its numerical advantages are stressed.

Moreover, the proposed method for choosing splines has the important degree of freedom regarding the possibility of approximating separately the semigroup governing the system and its adjoint. This allows us to use splines of any order [3]. This is very promising for obtaining very good performances in the future.

Future work will involve the infinite horizon LQG problem, which in this paper has been only sketched. For such a problem, the approximation scheme developed here can be used, and convergence of the type in paper [14] can be obtained. As a final remark, we would like to stress that the methodology presented in this paper can involve more than one type of approximation for the three equations governing the LQG stochastic control in order to get the best combination of theoretical and numerical convergences of approximation schemes developed until now [2, 3, 24, 26, 27, 40].

Appendix.

LEMMA A.1. *For any* $\boldsymbol{y}_0 \in \mathbb{R}^N$ *and for any function* $\boldsymbol{f} \in C^0([-r,0];\mathbb{R}^N)$*, there exists a unique left-continuous function* $\boldsymbol{y}_1 : [-r,0] \mapsto \mathbb{R}^N$ *such that*

$$\boldsymbol{y}_1 - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0, \boldsymbol{y}_1)\chi_{[-r,-r_j]} = \boldsymbol{f}, \tag{A.1}$$

where $\boldsymbol{k}_j$ *are functions defined in* (2.22).

Proof. In the case $\delta = 1$ the summation vanishes and the lemma is trivially true. In the case $\delta > 1$, consider (A.1) in time instants $-r_k$

$$\boldsymbol{y}_1(-r_k) - \sum_{j=1}^{k} \frac{\boldsymbol{y}_1(-r_j) - \boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{y}_0}{\delta - j + 1} = \boldsymbol{f}(-r_k), \quad k = 1, \dots, \delta - 1, \tag{A.2}$$

which can be put in matrix form as

$$\begin{bmatrix} \boldsymbol{y}_1(-r_1) \\ \vdots \\ \boldsymbol{y}_1(-r_{\delta-1}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}(-r_1) \\ \vdots \\ \boldsymbol{f}(-r_{\delta-1}) \end{bmatrix} + H_{\delta,2} \begin{bmatrix} \boldsymbol{y}_1(-r_1) \\ \vdots \\ \boldsymbol{y}_1(-r_{\delta-1}) \end{bmatrix} - H_{\delta,2} \begin{bmatrix} \boldsymbol{A}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{A}_{\delta-1}^{\mathrm{T}} \end{bmatrix} \boldsymbol{y}_0, \tag{A.3}$$

where matrix $H_{\delta,2}$ is defined as follows ($\boldsymbol{I}_N$ is the $N\times N$ identity matrix):

$$H_{\delta,2} = \begin{bmatrix} \frac{1}{\delta}\boldsymbol{I}_N & 0 & \cdots & 0 & 0 \\ \frac{1}{\delta}\boldsymbol{I}_N & \frac{1}{\delta-1}\boldsymbol{I}_N & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{\delta}\boldsymbol{I}_N & \frac{1}{\delta-1}\boldsymbol{I}_N & \cdots & \frac{1}{3}\boldsymbol{I}_N & 0 \\ \frac{1}{\delta}\boldsymbol{I}_N & \frac{1}{\delta-1}\boldsymbol{I}_N & \cdots & \frac{1}{3}\boldsymbol{I}_N & \frac{1}{2}\boldsymbol{I}_N \end{bmatrix}. \tag{A.4}$$

Now, let us define a vector $\boldsymbol{\eta}\in\mathbb{R}^{(\delta-1)N}$ as

$$\boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{\eta}_1 \\ \vdots \\ \boldsymbol{\eta}_{\delta-1} \end{bmatrix} = (\boldsymbol{I}_{(\delta-1)N} - H_{\delta,2})^{-1}\left(\begin{bmatrix} \boldsymbol{f}(-r_1) \\ \vdots \\ \boldsymbol{f}(-r_{\delta-1}) \end{bmatrix} - H_{\delta,2}\begin{bmatrix} \boldsymbol{A}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{A}_{\delta-1}^{\mathrm{T}} \end{bmatrix}\boldsymbol{y}_0 \right) \tag{A.5}$$

and the left-continuous function

$$\bar{\boldsymbol{y}}_1(\vartheta) = \begin{cases} \boldsymbol{\eta}_i, & \vartheta = -r_i, \\ \boldsymbol{f}(\vartheta) + \displaystyle\sum_{j=1}^{\delta-1}\frac{\boldsymbol{\eta}_j - \boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{y}_0}{\delta-j+1}\,\chi_{[-r,-r_j]}(\vartheta), & \vartheta \neq -r_i, \end{cases} \tag{A.6}$$

in which $i = 1,\dots,\delta-1$. It is readily verified that $\bar{\boldsymbol{y}}_1$ satisfies (A.1).

Uniqueness is proved by recognizing that any other function $\tilde{\boldsymbol{y}}_1$ satisfying (A.1) verifies also (A.3), and therefore $\tilde{\boldsymbol{y}}_1(-r_k) = \bar{\boldsymbol{y}}_1(-r_k)$, $k = 1,\dots,\delta-1$. The difference between expression (A.1) with $\boldsymbol{y}_1 = \bar{\boldsymbol{y}}_1$ and the same expression in which $\boldsymbol{y}_1 = \tilde{\boldsymbol{y}}_1$ is used gives $\bar{\boldsymbol{y}}_1 - \tilde{\boldsymbol{y}}_1 = 0$. This concludes the proof of uniqueness. ∎

Proof of Proposition 2.2. Only (2.20) and (2.21) require a little mathematics. The case $\delta = 1$ (summations in (2.20) and (2.21) vanish) is a standard result [24, 43]. For the case $\delta > 1$, let $L : \mathcal{D}(L) \mapsto \boldsymbol{M}_2$ be the operator defined as (see (2.20), (2.21))

$$L : \mathcal{D}(L) \mapsto \boldsymbol{M}_2,$$

$$L\begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{bmatrix} = \begin{bmatrix} \delta\,\boldsymbol{y}_1(0) + \boldsymbol{A}_0^{\mathrm{T}}\boldsymbol{y}_0 \\ \dfrac{1}{g}\boldsymbol{A}_{01}^{\mathrm{T}}\boldsymbol{y}_0 - \dfrac{d}{d\vartheta}\left(\boldsymbol{y}_1 - \displaystyle\sum_{j=1}^{\delta-1}\boldsymbol{k}_j(\boldsymbol{y}_0,\boldsymbol{y}_1)\chi_{[-r,-r_j]}\right) \end{bmatrix},$$

$$\mathcal{D}(L) = \left\{ \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{bmatrix} \;\middle|\; \begin{array}{c} \boldsymbol{y}_0 \in \mathbb{R}^N, \quad \boldsymbol{A}_\delta^{\mathrm{T}}\boldsymbol{y}_0 = \boldsymbol{y}_1(-r), \\ \left(\boldsymbol{y}_1 - \displaystyle\sum_{j=1}^{\delta-1}\boldsymbol{k}_j(\boldsymbol{y}_0,\boldsymbol{y}_1)\chi_{[-r,-r_j]}\right) \in W^{1,2} \end{array} \right\}.$$

We will show that

(a) for every $\boldsymbol{x}\in\mathcal{D}(\boldsymbol{A})$ and every $y\in\mathcal{D}(L)$, it is $(\boldsymbol{y},\boldsymbol{Ax}) - (L\boldsymbol{y},\boldsymbol{x}) = 0$;

(b) the set $\mathcal{D}(L)$ is dense in M_2.

These two items together state that L is the adjoint of $\boldsymbol{A}$, that is, $\boldsymbol{A}^*$ as defined in (2.20), (2.21).

Let us prove item (a). Take $\boldsymbol{x}\in\mathcal{D}(\boldsymbol{A})$ and $\boldsymbol{y}\in\mathcal{D}(L)$, let us show that $(\boldsymbol{y},\boldsymbol{Ax}) - (L\boldsymbol{y},\boldsymbol{x}) = 0$:

(A.7)
$$(\boldsymbol{y}, \boldsymbol{A}\boldsymbol{x}) - (L\boldsymbol{y}, \boldsymbol{x}) = \boldsymbol{y}_0^{\mathrm{T}} \sum_{j=1}^{\delta} \boldsymbol{A}_j \boldsymbol{x}_1(-r_j) - \delta\, \boldsymbol{y}_1^{\mathrm{T}}(0)\boldsymbol{x}_0 + \int_{-r}^{0} g(\vartheta) \left[\boldsymbol{y}_1^{\mathrm{T}}(\vartheta) \tfrac{d}{d\vartheta}\boldsymbol{x}_1(\vartheta) + \tfrac{d}{d\vartheta}\left(\boldsymbol{y}_1 - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0,\boldsymbol{y}_1)\chi_{[-r,-r_j]}\right)^{\mathrm{T}} \boldsymbol{x}_1(\vartheta)\right] d\vartheta.$$

$\boldsymbol{y}_1 - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0,\boldsymbol{y}_1)\chi_{[-r,-r_j]}$ being absolutely continuous, the integral term in (A.7) can be rewritten as

$$\begin{aligned}(A.8)\quad &\sum_{i=1}^{\delta} \int_{-r_i}^{-r_{i-1}} (\delta - i + 1) \frac{d}{d\vartheta}\left[\left(\boldsymbol{y}_1 - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0,\boldsymbol{y}_1)\chi_{[-r,-r_j]}\right)^{\mathrm{T}} \boldsymbol{x}_1(\vartheta)\right] d\vartheta \\ &+ \sum_{i=1}^{\delta} \int_{-r_i}^{-r_{i-1}} (\delta - i + 1) \left(\sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0,\boldsymbol{y}_1)\chi_{[-r,-r_j]}\right)^{\mathrm{T}} \frac{d}{d\vartheta}\boldsymbol{x}_1(\vartheta) d\vartheta\end{aligned}$$

and after a simple computation

$$\begin{aligned}(A.9)\quad &\sum_{i=1}^{\delta} (\delta - i + 1) \left[\left(\boldsymbol{y}_1 - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0,\boldsymbol{y}_1)\chi_{[-r,-r_j]}\right)^{\mathrm{T}} \boldsymbol{x}_1\right]_{-r_i}^{-r_{i-1}} \\ &+ \sum_{i=2}^{\delta} (\delta - i + 1) \left(\sum_{h=1}^{i-1} \boldsymbol{k}_h(\boldsymbol{y}_0,\boldsymbol{y}_1)\right)^{\mathrm{T}} [\boldsymbol{x}_1(-r_{i-1}) - \boldsymbol{x}_1(-r_i)] \\ &= \delta \boldsymbol{y}_1^{\mathrm{T}}(0)\, \boldsymbol{x}_1(0) - \boldsymbol{y}_1^{\mathrm{T}}(-r)\, \boldsymbol{x}_1(-r) - \sum_{i=1}^{\delta-1} \boldsymbol{y}_0^{\mathrm{T}} \boldsymbol{A}_i \boldsymbol{x}_1(-r_i).\end{aligned}$$

Now, replacing the integral term in (A.7) with the above expression and taking into account that $\boldsymbol{x}_1(0) = \boldsymbol{x}_0$ and $\boldsymbol{y}_1(-r) = \boldsymbol{A}_\delta^{\mathrm{T}} \boldsymbol{y}_0$, it follows that $(\boldsymbol{y}, \boldsymbol{A}\boldsymbol{x}) - (L\boldsymbol{y}, \boldsymbol{x}) = 0$.

Let us now prove item (b). It is sufficient to prove that $\mathcal{D}(L)$ is dense in $\mathbb{R}^N \times W^{1,2}$. Let $\boldsymbol{y}_0 \in \mathbb{R}^N$, $\boldsymbol{y}_1 \in W^{1,2}$. Consider the following sequence of functions $\boldsymbol{y}_{1,k} \in W^{1,2}$, defined for integers $k > \frac{1}{r - r_{\delta-1}}$:

(A.10)
$$\boldsymbol{y}_{1,k}(\vartheta) = \begin{cases} \boldsymbol{y}_1(\vartheta), & \vartheta \in [-r + \frac{1}{k}, 0], \\ (1 - k(\vartheta + r))\boldsymbol{A}_\delta^{\mathrm{T}} \boldsymbol{y}_0 + k(\vartheta + r)\boldsymbol{y}_1(-r + \frac{1}{k}), & \vartheta \in [-r, -r + \frac{1}{k}). \end{cases}$$

Being that $-r + 1/k < -r_{\delta-1}$, it is $\boldsymbol{y}_{1,k}(-r_j) = \boldsymbol{y}_1(-r_j)$ for $j = 1, \ldots, \delta - 1$. As $\boldsymbol{y}_1$ is uniformly bounded in $[-r, 0]$, given any positive ϵ there exists k_ϵ such that

$$\|\boldsymbol{y}_1 - \boldsymbol{y}_{1,k}\|_{L_2} < \frac{\epsilon}{2} \quad \text{for} \quad k > k_\epsilon. \tag{A.11}$$

Note that, since $\boldsymbol{y}_{1,k}(-r) = \boldsymbol{A}_d^{\mathrm{T}} \boldsymbol{y}_0$, if $\delta = 1$, then $\left[\begin{smallmatrix}\boldsymbol{y}_0 \\ \boldsymbol{y}_{1,k}\end{smallmatrix}\right] \in \mathcal{D}(\boldsymbol{A}^*)$, and the density of $\mathcal{D}(\boldsymbol{A}^*)$ in $\mathbb{R}^N \times W^{1,2}$ and hence in $\boldsymbol{M}_2$ is proved.

If $\delta > 1$ in general $\left[\begin{smallmatrix}\boldsymbol{y}_0 \\ \boldsymbol{y}_{1,k}\end{smallmatrix}\right] \notin \mathcal{D}(\boldsymbol{A}^*)$. For any integer $n > \sup_{j=1,2,\ldots,d} \frac{1}{r_j - r_{j-1}}$ it is convenient to define δ functions in $W^{1,2}$ as follows:

$$\chi_j^n(\vartheta) = \begin{cases} \chi_{[-r,-r_j]}(\vartheta), & \vartheta \notin (-r_j, -r_j + \frac{1}{n}), \\ -n(\vartheta + r_j - \frac{1}{n}), & \vartheta \in (-r_j, -r_j + \frac{1}{n}), \end{cases} \quad \text{for} \quad j = 1, \ldots, \delta - 1, \tag{A.12}$$

are such that

$$\|\chi_j^n - \chi_{[-r,-r_j]}\|_{L_2} = \frac{1}{\sqrt{3n}}, \quad \text{for } j = 1, \dots, \delta - 1. \tag{A.13}$$

By Lemma A.1, for any $n > \sup_{j=1,2,\dots,d} \frac{1}{r_j - r_{j-1}}$, there exists a function $\boldsymbol{y}_{1,k,n}$ such that

$$\boldsymbol{y}_{1,k,n} - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0, \boldsymbol{y}_{1,k,n})\chi_{[-r,-r_j]} = \boldsymbol{y}_{1,k} - \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0, \boldsymbol{y}_{1,k})\chi_j^n \tag{A.14}$$

(note that the right-hand side term is in C^0 and therefore Lemma A.1 can be applied).

It can be shown that $\boldsymbol{y}_{1,k,n}(-r_j) = \boldsymbol{y}_{1,k}(-r_j)$, $j = 1, \dots, \delta$, so that $\left[\begin{smallmatrix}\boldsymbol{y}_0 \\ \boldsymbol{y}_{1,k,n}\end{smallmatrix}\right] \in \mathcal{D}(\boldsymbol{A}^*)$.

Moreover,

$$\begin{aligned} \left\| \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_{1,k} \end{bmatrix} - \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_{1,k,n} \end{bmatrix} \right\|_{\boldsymbol{M}_2} &= \left\| \sum_{j=1}^{\delta-1} \boldsymbol{k}_j(\boldsymbol{y}_0, \boldsymbol{y}_1)(\chi_j^n - \chi_{[-r,-r_j]}) \right\|_{L_2} \\ &\le \sum_{j=1}^{\delta-1} \left\| \boldsymbol{k}_j(\boldsymbol{y}_0, \boldsymbol{y}_1) \right\| \frac{1}{\sqrt{3n}}, \end{aligned} \tag{A.15}$$

where formula (A.13) is used.

Thus, there exists an integer n_ϵ such that for $n > n_\epsilon$ it is

$$\left\| \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_{1,k} \end{bmatrix} - \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_{1,k,n} \end{bmatrix} \right\|_{\boldsymbol{M}_2} < \frac{\epsilon}{2}. \tag{A.16}$$

Finally, for any pair k, n such that $k > k_\epsilon$ and $n > n_\epsilon$ it is

$$\left\| \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{bmatrix} - \begin{bmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_{1,k,n} \end{bmatrix} \right\|_{\boldsymbol{M}_2} < \epsilon, \tag{A.17}$$

which proves the density of $\mathcal{D}(L)$ in $\boldsymbol{M}_2$. ∎

Remark A.2 The proof of this proposition concerning the adjoint operator $\boldsymbol{A}^\star$ can also be done by methodology shown in [15]. Using standard Lax–Milgram-type representation theorems, a relationship follows between equivalent inner products, so that an adjoint operator in a given inner product can be found by another one obtained in an equivalent inner product (see [15] and references therein).

LEMMA A.3. *For any nonnegative λ the matrix $H_p(\lambda)$ defined in* (4.59) *is nonsingular.*

Proof. The expression of $H_p(\lambda)$ is here reported for the reader's convenience:

$$H_p(\lambda) = \boldsymbol{I}_{(\delta-1)N} - H_{\delta,2} + \begin{bmatrix} \boldsymbol{I}_N e^{-\lambda(r-r_1)} \\ \vdots \\ \boldsymbol{I}_N e^{-\lambda(r-r_{\delta-1})} \end{bmatrix} h_{\delta,2}. \tag{A.18}$$

As a first step, nonsingularity of matrix $\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2}$ is proved. By the definition of $H_{\delta,2}$ in (A.4) it is

(A.19)

$$\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2} = \begin{bmatrix} \boldsymbol{I}_N - \frac{1}{\delta}\boldsymbol{I}_N & 0 & \cdots & 0 & 0 \\ -\frac{1}{\delta}\boldsymbol{I}_N & \boldsymbol{I}_N - \frac{1}{\delta-1}\boldsymbol{I}_N & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{1}{\delta}\boldsymbol{I}_N & -\frac{1}{\delta-1}\boldsymbol{I}_N & \cdots & \boldsymbol{I}_N - \frac{1}{3}\boldsymbol{I}_N & 0 \\ -\frac{1}{\delta}\boldsymbol{I}_N & -\frac{1}{\delta-1}\boldsymbol{I}_N & \cdots & -\frac{1}{3}\boldsymbol{I}_N & \boldsymbol{I}_N - \frac{1}{2}\boldsymbol{I}_N \end{bmatrix}.$$

A direct computation shows that the inverse of $\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2}$ is

$$(\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2})^{-1} = \boldsymbol{I}_{N(\delta-1)} + \overline{H}_{\delta,2}, \tag{A.20}$$

where matrix $\overline{H}_{\delta,2}$ is defined as

$$\overline{H}_{\delta,2} = \begin{bmatrix} \frac{1}{\delta-1}\boldsymbol{I}_N & 0 & \cdots & 0 & 0 \\ \frac{1}{\delta-2}\boldsymbol{I}_N & \frac{1}{\delta-2}\boldsymbol{I}_N & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{2}\boldsymbol{I}_N & \frac{1}{2}\boldsymbol{I}_N & \cdots & \frac{1}{2}\boldsymbol{I}_N & 0 \\ \boldsymbol{I}_N & \boldsymbol{I}_N & \cdots & \boldsymbol{I}_N & \boldsymbol{I}_N \end{bmatrix}. \tag{A.21}$$

The verification can be made writing the following expression for the kth column block of matrix $\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2}$ as

$$\begin{bmatrix} 0 \\ \vdots \\ \boldsymbol{I}_N \\ \vdots \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ \vdots \\ \frac{1}{\delta-(k-1)}\boldsymbol{I}_N \\ \vdots \\ \frac{1}{\delta-(k-1)}\boldsymbol{I}_N \end{bmatrix} \tag{A.22}$$

(the first $k-1$ blocks are zero) and the following expression for the jth row block of matrix $\left(\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2}\right)^{-1}$ as

$$[0 \ \cdots \ 0 \ \ \boldsymbol{I}_N \ \ 0 \ \cdots 0] + \left[\tfrac{1}{\delta-j}\boldsymbol{I}_N \ \ \ldots \ \ \tfrac{1}{\delta-j}\boldsymbol{I}_N \ \ 0 \ \cdots \ 0 \right] \tag{A.23}$$

(the first j blocks are nonzero). The product when $j < k$ is a sum of zeroes and is trivially zero. It can also be verified that when $j > k$, the product gives zero, and when $j = k$, the product gives $\boldsymbol{I}_N$. This verifies the expression (A.20) for $(\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2})^{-1}$.

Given the invertible matrix $\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2}$, the determinant of $H_p(\lambda)$ can be written as follows:

$$\det(\mathbf{I}_{N(\delta-1)} - H_{\delta,2}) \cdot \det\left(\mathbf{I}_{(\delta-1)N} + (\mathbf{I}_{N(\delta-1)} - H_{\delta,2})^{-1} \begin{bmatrix} \mathbf{I}_N e^{-\lambda(r-r_1)} \\ \vdots \\ \mathbf{I}_N e^{-\lambda(r-r_{\delta-1})} \end{bmatrix} h_{\delta,2} \right). \tag{A.24}$$

Since for any pair of matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times n}$

$$\det(\boldsymbol{I}_n + A \cdot B) = \det(\boldsymbol{I}_m + B \cdot A), \tag{A.25}$$

the determinant of $H_p(\lambda)$ can also be written as

$$\det(\mathbf{I}_{N(\delta-1)} - H_{\delta,2}) \det\left(\mathbf{I}_{(\delta-1)B} + h_{\delta,2}(\mathbf{I}_{N(\delta-1)} - H_{\delta,2})^{-1} \begin{bmatrix} \mathbf{I}_N e^{-\lambda(r-r_1)} \\ \vdots \\ \mathbf{I}_N e^{-\lambda(r-r_{\delta-1})} \end{bmatrix} \right). \tag{A.26}$$

Recalling the expression of $h_{\delta,2}$ defined in (4.56) it follows that

$$h_{\delta,2}\left(\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2}\right)^{-1} \begin{bmatrix} \boldsymbol{I}_N e^{-\lambda(r-r_1)} \\ \vdots \\ \boldsymbol{I}_N e^{-\lambda(r-r_{\delta-1})} \end{bmatrix} = c(\lambda)\boldsymbol{I}_N, \tag{A.27}$$

where $c(\lambda)$ is the sum of positive terms that are functions of λ. Therefore it is

$$\det(H_p(\lambda)) = \det(\boldsymbol{I}_{N(\delta-1)} - H_{\delta,2}) \det((1+c(\lambda))\boldsymbol{I}_N) = \frac{1}{\delta^N}(1+c(\lambda))^N, \tag{A.28}$$

and this proves the nonsingularity of $H_p(\lambda)$ for any nonnegative λ.

REFERENCES

[1] A. V. Balakrishnan, *Applied Functional Analysis*, Springer-Verlag, New York, 1981.

[2] H. T. Banks and J. A. Burns, *Hereditary control problems: Numerical methods based on averaging approximations*, SIAM J. Control Optimization, 16 (1978), pp. 169–208.

[3] H. T. Banks and F. Kappel, *Spline approximations for functional differential equations*, J. Differential Equations, 34 (1979), pp. 496–522.

[4] H. T. Banks, G. I. Rosen, and K. Ito, *A spline based technique for computing Riccati operators and feedback controls in regulator problems for delay equations*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 830–855.

[5] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter, *Representation and Control of Infinite Dimensional Systems*, Birkhaüser, Boston, 1992.

[6] J. Burns, K. Ito, and G. Propst, *On nonconvergence of adjoint semigroups for control systems with delays*, SIAM J. Control Optim., 26 (1988), pp. 1441–1454.

[7] F. Colonius, *Stable and regular reachability for relaxed hereditary differential systems*, SIAM J. Control Optim., 20 (1982), pp. 675–694.

[8] F. Colonius, *Addendum: Stable and regular reachability of relaxed hereditary differential systems*, SIAM J. Control Optim., 23 (1985), pp. 803–807.

[9] F. Colonius, *The maximum principle for relaxed hereditary differential systems with function space end condition*, SIAM J. Control Optim., 5 (1982), pp. 695–712.

[10] R. F. Curtain and A. J. Pritchard, *Infinite Dimensional Linear Systems Theory*, Springer-Verlag, Berlin, 1978.

[11] M. C. Delfour, *The linear quadratic optimal control problem for hereditary differential systems: Theory and numerical solution*, Appl. Math. Optim., 3 (1977), pp. 101–162.

[12] M. C. Delfour, *The linear-quadratic optimal control problem with delays in state and control variables: A state space approach*, SIAM J. Control Optim., 24 (1986), pp. 835–883.

[13] M. C. Delfour and S. K. Mitter, *Controllability, observability and optimal feedback control of affine hereditary differential systems*, SIAM J. Control, 10 (1972), pp. 298–328.

[14] A. De Santis, A. Germani, and L. Jetto, *Approximation of the algebraic Riccati equation in the Hilbert space of Hilbert-Schmidt operators*, SIAM J. Control Optim., 4 (1993), pp. 847–874.

[15] R. H. Fabiano, *Stability preserving spline approximations for scalar functional differential equations*, Computers Math. Appl., 8 (1995), pp. 87–94.

[16] R. H. Fabiano and K. Ito, *Semigroup theory and numerical approximation for equations in linear viscoelasticity*, SIAM J. Math. Anal., 21 (1990), pp. 374–393.

[17] A. Feliachi and A. Thowsen, *Memoryless stabilization of linear delay differential systems*, IEEE Trans. Automat. Control, 2 (1981), pp. 586–587.

[18] A. Germani and L. Jetto, *A twofold spline scheme approximation method for the LQG control problem on Hilbert spaces*, in Proceedings of the 12th IFAC Triennal World Congress, Sydney, Australia, 1993, Vol. 2, pp. 401–404.

[19] A. Germani, L. Jetto, C. Manes, and P. Pepe, *The LQG control problem for a class of hereditary systems: A method for computing its approximate solution*, in Proceedings of the 33rd IEEE Conference on Decision and Control, Orlando, FL, 1993, Vol. 2, IEEE, New York, 1993, pp. 1362–1367.

[20] A. Germani, L. Jetto, and M. Piccioni, *Galerkin approximation for optimal filtering of infinite-dimensional linear systems*, SIAM J. Control Optim., 26 (1988), pp. 1287–1305.

[21] A. Germani, C. Manes, and P. Pepe, *Numerical solution for optimal regulation of stochastic hereditary systems with multiple discrete delays*, in Proceedings of the 34th IEEE Conference on Decision and Control, New Orleans, LA, 1995, Vol. 2, IEEE, New York, pp. 1497–1502.

[22] A. Germani, C. Manes, and P. Pepe, *Implementation of an LQG control scheme for linear systems with delayed feedback action*, in Proceedings of the 3rd European Control Conference, Roma, Italy, 1995, Vol. 4, EUCA, Rome, 1995, pp. 2886–2891.

[23] J. S. Gibson, *The Riccati integral equations for optimal control problems on Hilbert spaces*, SIAM J. Control Optim., 17 (1979), pp. 537–565.

[24] J. S. Gibson, *Linear-quadratic optimal control of hereditary differential systems: Infinite-dimensional Riccati equations and numerical approximations*, SIAM J. Control Optim., 21 (1983), pp. 95–139.

[25] K. Ito, *Finite-dimensional compensators for infinite-dimensional systems via Galerkin-type approximation*, SIAM J. Control Optim., 28 (1990), pp. 1251–1269.

[26] K. Ito and F. Kappel, *A uniformly differentiable approximation scheme for delay systems using splines*, Appl. Math. Optim., 23 (1991), pp. 217–262.

[27] F. Kappel and D. Salamon, *Spline approximation for retarded systems and the Riccati equation*, SIAM J. Control Optim., 25 (1987), pp. 1082–1117.

[28] F. Kappel and D. Salamon, *An approximation theorem for the algebraic Riccati equation*, SIAM J. Control Optim., 28 (1990), pp. 1136–1147.

[29] F. Kappel and D. Salamon, *On the stability properties of spline approximations for retarded systems*, SIAM J. Control Optim., 27 (1989), pp. 407–431.

[30] K. Kunisch, *Approximation schemes for the linear-quadratic optimal control problem associated with delay equations*, SIAM J. Control Optim., 20 (1982), pp. 506–540.

[31] R. H. Kwong, *A stability theory for the linear-quadratic-Gaussian problem for systems with delays in the state, control, and observations*, SIAM J. Control Optim., 18 (1980), pp. 49–75.

[32] I. Lasiecka and A. Manitius, *Differentiability and convergence rates of approximating semigroups for retarded functional differential equations*, SIAM J. Numer. Anal, 25 (1988), pp. 883–907.

[33] A. Manitius and H. T. Tran, *Numerical approximations for hereditary systems with input and output delays: Convergence results and convergence rates*, SIAM J. Control Optim., 32 (1994), pp. 1332–1363.

[34] T. Mori, *Criteria for asymptotic stability of linear time delay systems*, IEEE Trans. Automat. Control, 2 (1985), pp. 158–161.

[35] K. A. Morris, *Convergence of controllers designed using state-space techniques*, IEEE Trans. Automat. Control, 10 (1994), pp. 2100–2104.

[36] K. A. Morris, *Design of finite-dimensional controllers for infinite-dimensional systems by approximation*, J. Math. Systems Estim. Control, 4 (1994), pp. 1–30.

[37] L. Pandolfi, *The standard regulator problem for systems with input delays. An approach through singular control theory*, Appl. Math. Optim., 31 (1995), pp. 119–136.

[38] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.

[39] P. Pepe, *Il Controllo LQG dei Sistemi con Ritardo*, Ph. D. Thesis, Department of Electrical Engineering, Università degli studi dell'Aquila, Italy, 1996.

[40] G. Propst, *Piecewise linear approximation for hereditary control problems*, SIAM J. Control Optim., 28 (1990), pp. 70–96.

[41] D. Salamon, *Structure and stability of finite-dimensional approximations for functional-differential equations*, SIAM J. Control Optim., 23 (1985), pp. 928–951.

[42] M. H. Schultz, *Spline Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[43] R. B. Vinter, *On the evolution of the state of linear differential delay equations in M^2: Properties of the generator*, J. Inst. Math. Appl., 21 (1978), pp. 13–23.